REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	15.Sep.03	DISSERTATION

4. TITLE AND SUBTITLE
"VALIDATION OF A CLINICAL PREDICTION RULE TO IDENTIGY PATIENTS WITH LOW BACK PAIN LIKELY TO BENEFIT FROM SPINAL MANIPULATION"

5. FUNDING NUMBERS

6. AUTHOR(S)
CAPT CHILDS JOHN D

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF PITTSBURGH

8. PERFORMING ORGANIZATION REPORT NUMBER
CI02-1280

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

20030923 110

14. SUBJECT TERMS

15. NUMBER OF PAGES
249

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

THE VIEWS EXPRESSED IN THIS ARTICLE ARE THOSE OF THE AUTHOR AND DO NOT REFLECT THE OFFICIAL POLICY OR POSITION OF THE UNITED STATES AIR FORCE, DEPARTMENT OF DEFENSE, OR THE U.S. GOVERNMENT

VALIDATION OF A CLINICAL PREDICTION RULE TO IDENTIFY PATIENTS LIKELY TO BENEFIT FROM SPINAL MANIPULATION: A RANDOMIZED CLINICAL TRIAL

by

John David Childs

BS, Biology, U.S. Air Force Academy, 1994

MPT, Physical Therapy, U.S. Army-Baylor University, 1996

MBA, Business, University of Arizona, 2000

MS, Musculoskeletal Physical Therapy, University of Pittsburgh, 2002

Submitted to the Graduate Faculty of the

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2003

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

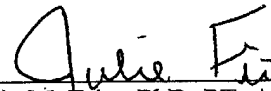
by

John D. Childs

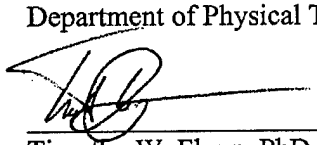
It was defended on

June 24th, 2003

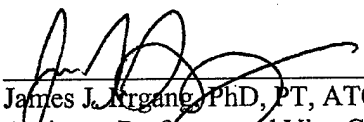
and approved by



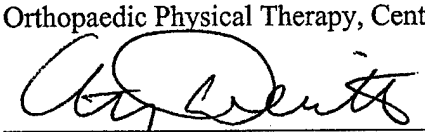
Julie M. Fritz, PhD, PT, ATC (Dissertation Chair)
Assistant Professor
Department of Physical Therapy, University of Pittsburgh



Timothy W. Flynn, PhD, PT, OCS, FAAOMPT
Program Director and Associate Professor,
US Army-Baylor University Graduate Program in Physical Therapy



James J. Irrgang, PhD, PT, ATC
Assistant Professor and Vice Chairman for Clinical Services,
Department of Physical Therapy, University of Pittsburgh and
Vice President of QI and Outcomes and Director of Sports and
Orthopaedic Physical Therapy, Centers for Rehab Services



Anthony Delitto, PhD, PT, FAPTA
Associate Professor and Chair
Department of Physical Therapy, University of Pittsburgh

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

John D. Childs

It was defended on

June 24th, 2003

and approved by

Julie M. Fritz, PhD, PT, ATC (Dissertation Chair)
Assistant Professor
Department of Physical Therapy, University of Pittsburgh

Timothy W. Flynn, PhD, PT, OCS, FAAOMPT
Program Director and Associate Professor,
US Army-Baylor University Graduate Program in Physical Therapy

James J. Irrgang, PhD, PT, ATC
Assistant Professor and Vice Chairman for Clinical Services,
Department of Physical Therapy, University of Pittsburgh and
Vice President of QI and Outcomes and Director of Sports and
Orthopaedic Physical Therapy, Centers for Rehab Services

Anthony Delitto, PhD, PT, FAPTA
Associate Professor and Chair
Department of Physical Therapy, University of Pittsburgh

VALIDATION OF A CLINICAL PREDICTION RULE TO IDENTIFY PATIENTS LIKELY TO BENEFIT FROM SPINAL MANIPULATION: A RANDOMIZED CLINICAL TRIAL

John D. Childs, PhD

University of Pittsburgh, 2003

Purpose: The primary aim of this study was to validate a clinical prediction rule (CPR) to identify patients with low back pain (LBP) likely to benefit from spinal manipulation. **Subjects:** 131 consecutive patients referred for physical therapy. Patients with positive neurologic signs or other red flags for spinal manipulation were excluded. **Method:** A multicenter, randomized clinical trial. After completing a standardized history and physical examination, patients were randomly assigned to receive spinal manipulation (n=70) or a stabilization exercise intervention (n=61). Patients were seen in physical therapy twice the first week, then once a week for the next three weeks, for a total of five sessions. A single manipulative intervention was used for patients who received spinal manipulation during each of the first two sessions, who then completed the stabilization exercise intervention for the remaining three weeks. Patients who achieved at least a 50% improvement in their Oswestry Disability Questionnaire (ODQ) score were classified as a success. Patients who met at least 4/5 criteria in the CPR were classified as positive. **Analyses:** A 2*2*3 repeated measures multivariate analysis of variance (MANOVA) was performed, followed by a Bonferroni procedure for planned comparisons. Sensitivity, specificity, and positive and negative likelihood ratios (LR) with associated 95% confidence intervals were calculated. **Results:** There was a significant three-way CPR*Intervention*Time interaction for the overall repeated measures MANOVA ($p<.001$). Patients classified as positive on the CPR and received spinal manipulation achieved 2.5 times the minimum clinically important difference

(MCID) on the ODQ compared to patients classified as negative on the CPR and received spinal manipulation and 3.4 times the MCID compared to patients classified as positive on the CPR but received the stabilization exercise intervention ($p<.001$). These results were maintained at the four-week follow-up ($p<.003$). With a positive LR of 13.2 (3.4, 52.1) and based on a pre-test probability of success of 44.3%, this translates into a post-test probability of success of 91.2%.

Conclusions: The results of this study support the validity of the spinal manipulation CPR.

Clinical Relevance: Clinicians can accurately identify patients with LBP likely to benefit from spinal manipulation.

ACKNOWLEDGMENTS

Thank you to our fearless leader and department chair, Tony Delitto. Tony taught me how to stay out of people's way and let them excel, just sticking your head in as needed to tear down the "red tape" that often stands in the way of progress. There is no better mentor or leader in our profession. It has been a privilege to be a small part of the Pittsburgh family. Many opportunities have been placed in my path simply based on being associated with many of the best in our profession.

Thank you to my dissertation chair and adviser, Julie Fritz, for your being more than a mentor, but a genuine friend who works harder than almost anyone I know. Thank you for giving me the opportunity to run with a great project. I'm looking forward to years or friendship and working together.

Thank you to my good friend and clinical mentor, Dick Erhard, who taught me much about how to treat patients and a little about everything else in life. There's not much this man doesn't know. By the way, now when is it that you are going to retire?

Thank you to Jay Irrgang and Kelly Fitzgerald for introducing me to the PhD-level research process. You both have been first class mentors. Jay, you are a no-sleep machine. Kelly, you taught me how to never give up on a research question that deserves an answer.

Thank you to my first clinical mentor and boss, Kevin Johnson, who gave me countless opportunities to grow professionally. Although Tim Flynn recently coined the phrase, Kevin was the first to show me to just "manip, and move on". He is also a subject recruiting machine.

Thank you to all of the physical therapists who assisted me with data collection. I could not have been successful without you: Kevin Johnson, Guy Majkowski, Maria West, Evan Kelley, David Browder, Mike Blowers, Sherri Morrow, Brian Langford, Jeff McGuire, Cory Middel, Trevor Petrou, and Kelly Rhoden.

Thank you to one of my best friends and doctoral colleague, Sara Piva, who is almost as unable as I am to say "No" to a new project or adventure. This experience would not have been nearly as enjoyable if it weren't for your friendship. I look forward to years of working together.

Thank you to my friends and physical therapy colleagues from the military. To Rob Wainner, who, upon my being selected by the Air Force to pursue a PhD, let me know that Pitt was really the only program to consider. Your counsel was truly wise. To Tim Flynn, the "recovering biomechanist", it has been a privilege having you my committee. You are one of the few who is a master clinician and researcher, all in one package. Julie Whitman, you've been a great friend and colleague with energy that doesn't stop. Andrew Bennett, as a very young therapist, you've raised the bar for clinical excellence, showing me just how good a clinician can be.

At the risk of forgetting someone, I'd like to thank a number of faculty, former doctoral colleagues, and other physical therapists that I've had the privilege of working with and getting

to know at the University of Pittsburgh. You each have made a unique impact on my life: Steve George, Greg Hicks, Michelle Vignovic, Tara Ridge, Erica Baum, Joel Robb, Christian Lyons, David Browder, Pat Sparto, Jen Brach, and Kathy Kelly.

Thank you to each of my parents for their unique contribution to my being able to succeed, but especially to Mom, who set an example on how to pursue excellence in everything you do.

Thank you to my wonderful children, who, although didn't fully understand what Daddy was up to, patiently let me work from home. They had to listen to me say too many times, "I'll be with you in just a little while." I look forward to saying that much less often. Their rarely being impressed with anything I might do professionally helps me realize what really matters most in life.

Thank you to my beautiful wife You sacrificed countless hours of your own time letting me pursue my dreams. You made a decision early after our graduation from the Academy that becoming a fighter pilot had the potential to place an excessive burden on our family. Just as you already are as my wife and Mommy to our girls, no one doubts that you would have been among the "best of the best". Many call it a sacrifice, but you see it simply as a blessing. Only our children and I can truly appreciate the impact of your decision in the short run. However, the passage of time through the legacy you leave will uncover you selflessness and dedication to our family for generations to come. I love you more than you will ever know.

And lastly and most importantly, thank you to my Lord and Savior Jesus Christ for the provision of an opportunity to accomplish a dream.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Statement of the Problem

Because of the substantial impact of low back pain (LBP) on healthcare systems throughout the world, there is a need to identify effective interventions to reduce the disability associated with LBP. However, few interventions for patients with LBP have evidence for their effectiveness. One explanation for the failure of research to identify effective interventions is the inability to classify patients with LBP into more homogenous subgroups that are likely to likely benefit from a specific intervention. Spinal manipulation is one intervention that has some evidence for its effectiveness; however, the results are sometimes contradictory. Some studies have demonstrated spinal manipulation to be effective[1-7] while others have shown that it is not.[8-11] This confusion makes it difficult for clinicians to determine when manipulation should be used in the treatment of patients with LBP. A previous study conducted the first step in the development of a clinical prediction rule (CPR) to identify patients with LBP most likely to likely benefit from spinal manipulation.[12] Patients who met at least 4/5 criteria in the CPR improved their chances of success with spinal manipulation from 45% to 95%. Success was defined as achieving at least a 50% improvement on the Oswestry Disability Questionnaire (ODQ).

Although the results of this initial step in the development of a CPR are exciting, subsequent studies are necessary to validate the initial findings in a different patient population. This study was a multicenter randomized clinical trial (RCT) designed to validate the CPR in a different sample to identify patients with LBP likely to benefit from spinal manipulation. If the results of the initial study can be validated, clinicians will benefit by having an easy-to-use CPR to assist them in decision-making to identify patients likely to benefit from this intervention. Future

clinical trials can then be developed to test the implementation of the CPR in clinical practice on practice patterns, outcomes of care, and costs.

2. Background and Significance

2.1 Prevalence and Costs Associated with LBP

Next to the common cold, the complaint of back pain is the most common reason individuals visit a physician's office[13] and affects 60-90% of individuals during their lifetime.[14] The complaint of acute LBP alone was the fifth most common reason for individuals to seek physician assistance in 1995.[15] Billions of dollars in medical expenditures related to the treatment of LBP are incurred each year.[16,17] Because of the substantial impact of LBP on healthcare systems throughout the world, there is a need to identify effective interventions and prevention strategies.

2.2 Attempts to Identify Effective Interventions for Low Back Pain

Unfortunately, many interventions currently used in the management of patients with LBP do not have evidence to support their use, and attempts to identify effective interventions have often been futile.[18-21] Little evidence, if any, exists to support several interventions utilized by therapists, to include traction, biofeedback, heat and cold modalities, and transcutaneous electrical nerve stimulation (TENS).[20,21] Research that has assessed the effectiveness of exercise in patients with LBP also has been equivocal.[18,19] The evidence seems to advocate exercise for patients with chronic LBP but seems to be ineffective for patients with acute LBP.[18,19,22] It is possible that one explanation for the lack of evidence for many interventions for patients with LBP is the inability of researchers to define relevant subgroups of patients who are most likely to

benefit from the intervention being studied.[23,24] The inclusion criteria that used in research on interventions for the management of LBP are frequently too broad, including patients for whom the intervention would not be expected to be beneficial. Without the ability to match patients to specific interventions, clinicians are left without evidence for decision-making to select interventions for a particular patient.

2.3 Lumbopelvic Region Dysfunction a Potential Subgroup of Low Back Pain

Despite the controversy over the exact prevalence, research indicates that the lumbopelvic region, specifically the sacroiliac (SI) joint, is a possible source of LBP.[25,26] The term "SI joint dysfunction" is used to explain pain from a SI joint that exhibits no demonstrable lesion but is presumed to have some type of biomechanical disorder that causes that pain.[25] The biomechanical disorder may be a state of relative hypomobility within a portion of the joint's range of motion with subsequent altered positional relationships between the sacrum and ilium.[27] However, based on the lack of evidence to implicate the joint itself as the primary source of pain and dysfunction,[27-30] the term "lumbopelvic region" will be used hereafter to characterize this subgroup of patients.

2.3.1 The Traditional Approach for the Diagnosis of Lumbopelvic Region Dysfunction

Traditionally, clinicians have primarily utilized a pathoanatomical approach by relying on multiple clinical diagnostic tests purported to identify and diagnose dysfunction in the lumbopelvic region. Components of the physical examination commonly used in patients with suspected dysfunction in this area include tests designed to assess the symmetry of bony landmarks in the static position (static symmetry tests), tests to assess the symmetry of bony landmarks with movement (movement symmetry tests), and tests to reproduce symptoms

(provocation tests). The results of these tests are then typically used to guide the choice of a manual therapy intervention specific to the biomechanical disorder that is identified. Many tests are described in the peer-reviewed literature as being useful to identify dysfunction in the lumbopelvic region.[31-34]

2.3.2 Limitations of Lumbopelvic Region Diagnostic Tests for Clinical Decision-making

The utility of diagnostic tests to assess the lumbopelvic region for decision-making may be limited for several reasons. First, these tests have not been demonstrated to be sufficiently reliable to justify their use.[32-34] Secondly, little research has been conducted to establish the validity of these tests.[35] The few studies that have attempted to assess the validity for using these tests[27-30] largely refute the theoretical foundation for bony movement at the SI joint, a foundation upon which many of these tests are based. The poor reliability and validity for these tests contribute to clinicians frequently obtaining results that are contradictory among patients in which these tests are purported to be useful. Based on the overwhelming evidence that suggests these tests are not sufficiently reliable or valid, their use for decision-making appears to be limited. In fact, it has recently been suggested[36,37] that clinicians as a whole abandon these tests in pursuit of more reliable and clinically useful means by which to diagnose and treat dysfunction in the lumbopelvic region.

2.3.2.1 Reliability of Diagnostic Tests for Lumbopelvic Region Dysfunction

Reliability studies for diagnostic tests to assess dysfunction in the lumbopelvic region can be divided into those that have demonstrated acceptable reliability and studies that have failed to demonstrate acceptable reliability. Only a few studies that have assessed the reliability of these tests report acceptable reliability.[31,38-41] For those studies that have reported acceptable

reliability, limitations in the methodology exist such as using asymptomatic subjects,[38,39] tester awareness of the expected measure,[39] and unclear descriptions of the population and outcome measure.[40] Each of these limitations hinders the clinical utility of these tests. The only tests for which there appear to be an acceptable level of reliability using appropriate methodological procedures is the palpation of iliac crest height in sitting[41] and palpation of iliac crest height in standing.[31] However, little research has been conducted to establish the validity for any of these tests.

Most studies[25,32,35,40,42-53] have almost uniformly demonstrated poor reliability for these tests. Studies that have failed to demonstrate acceptable reliability for individual tests[45,46,50,54] have also suffered from methodological flaws such as lack of scientific rigor in the statistical methodology,[46] use of the non weight-bearing position,[51] failure to use the dominant eye to sight the landmarks, lack of sufficient distance between the palpating fingers and the evaluator's eyes,[45,46,50,54] subject fatigue after long standing periods,[50] and use of non-representative groups of subjects.[50]

O'Haire[51] suggested that the poor reliability might be partially attributed to errors in the precise location of the anatomical landmarks used in these tests. It has been proposed that better standardization of measures and the use of appropriate statistical methods in studies of dysfunction in the lumbopelvic region may result in improved measurement accuracy.[55] However, despite diligent efforts to standardize and refine the operational definitions and testing procedures associated with these tests, these efforts have largely not improved the reliability of these tests.

2.3.2.2 Validity of Diagnostic Tests for Lumbopelvic Region Dysfunction

Not only have diagnostic tests for lumbopelvic region dysfunction largely been demonstrated to be unreliable, but there is also little evidence to support their validity.[35] Clinicians who use these tests to guide decision-making in the selection of an intervention such as mobilization/manipulation for patients with LBP rarely use the results of a single test in isolation. Rather, they use the results of these tests in combination with other pertinent information from the patient's history and physical examination first to determine if a patient is appropriate for the intervention and if so, to help guide decision-making in the selection of the most appropriate technique. However, studies that have assessed the reliability of these tests tend not to consider other potentially relevant variables from the history and physical examination that may be used for decision-making.

Secondly, the use of these tests to guide treatment is largely based on theoretical principles that are not necessarily supported by the current scientific evidence. There is growing evidence that movement between the sacrum and the ilium is less than 2 mm and less than 2°,[27-29] an amount of motion so small that it may likely not even be detectable with palpation. Gerlach and Lierse[56] suggest that such a small amount of movement substantiates the importance of the ligamentous apparatus holding these structures together. Perhaps even more importantly, Tullberg et al[30] found using roentgen stereophotogrammetric that manipulation did not alter the position of the SI joint. It has been suggested that dysfunction in the lumbopelvic region could be caused by a slight joint derangement that would greatly alter the transmission of forces through the pelvis and low back area, thus representing a potential source of ongoing discomfort.[42,56] However, these findings largely fail to support the theoretical foundation for bony movement, a foundation upon

which many of these tests are based. Therefore, even if these tests could be reliably measured, the theoretical foundation may be seriously flawed, and the construct validity in their use for decision-making is lacking. Interestingly, recent evidence does suggest that manipulation increases gapping of the facet joints in the lumbar spine.[57]

Based on our experience and the suggestion of others,[30] it seems reasonable to suspect that lumbopelvic asymmetries in patients with LBP that may be addressed with manipulation may instead be attributed to changes in soft tissues in this region; however, little research has been conducted to examine the plausibility of this hypothesis. It is possible that discrepancy in side-to-side weight-bearing between the lower extremities could be a manifestation of soft tissue or other biomechanical dysfunction in the lumbopelvic region. Childs et al[58] found that patients with LBP demonstrated increased side-to-side weight-bearing asymmetry compared to healthy control subjects without LBP. Higher magnitudes of asymmetry were associated with increased levels of pain. A subsequent follow-on study[59] demonstrated significant improvements in iliac crest and weight-bearing asymmetry immediately after spinal manipulation, changes which were related to improvements in the patient's self-reported level of pain. Although the results of these studies do not lend direct evidence to support the validity of the role of soft tissues in their contribution to LBP, they do provide preliminary evidence to support the theoretical rationale for how manipulation may work to improve pain and function in these patients. Future work needs to be conducted to further examine the mechanisms through which spinal manipulation acts to improve pain and function.

Some have suggested that perhaps the results of a combination of tests may be more useful than the results of a single test in isolation.[25,44,60-62] Several studies[25,44,60,61] have attempted to determine the diagnostic validity of these tests in isolation and in combination using short-term pain relief after injection of the lumbopelvic region as the reference criterion to identify which patients exhibited dysfunction in the lumbopelvic region. None of these studies support using the results of these tests in isolation, and only one of the studies[60] supports using the results of a combination of tests. However, the validity of using injection as the reference criterion to identify which patients have dysfunction in the lumbopelvic region has not been demonstrated[63,63,64,64,65] and is not useful clinically. Only one other study has attempted to determine if the usefulness of these tests in combination is more meaningful than their use in isolation. Cibulka et al[62] described a cluster of four symmetry and movement tests (three of which must be positive to indicate an overall positive test) to identify patients with dysfunction in the lumbopelvic region and reported good reliability and validity. However, they used the presence or absence of LBP as the reference criterion to establish that dysfunction in the lumbopelvic region was in fact present.[62] This reference criterion assumes that dysfunction in the lumbopelvic region contributes to all cases of LBP, which does not make intuitive sense clinically. Perhaps more importantly, the sample included asymptomatic patients without LBP. For the results of a study assessing the validity of a diagnostic test to be useful, the sample should include an appropriate spectrum of patients in which clinicians would actually perform the tests. Clinicians do not routinely perform diagnostic tests in asymptomatic patients, thus they only useful to the extent they can distinguish between patients with the diagnosis of interest and other competing diagnoses that need to be ruled in or ruled out. Based on these factors, the

results of this study are limited. Additionally, a recent study[52] failed to replicate the findings of Cibulka et al[62] because the individual tests in the cluster were largely unreliable.

2.3.2.3 Lumbopelvic Region Diagnostic Test Results Often Conflicting

Based on our clinical experience, interpreting the results of diagnostic tests for dysfunction in the lumbopelvic region can frequently lead to conflicting findings, which may largely be attributed to their overall poor reliability and validity. Two case reports were recently published illustrating this phenomenon.[66] Using the traditional approach to examine a patient with suspected dysfunction in the lumbopelvic region, several of the classic diagnostic tests were performed. Based on the static tests in which the bony landmarks were palpated, one of the patients exhibited an apparent high PSIS and IC in conjunction with a low ASIS on the right in static standing. With the traditional approach, this combination of findings suggests the presence of an anteriorly rotated innominate on the right. The supine long-sitting test and the prone knee flexion test were performed next. The supine long-sitting test is a test that compares apparent leg lengths in the supine and long-sitting positions. To perform this test, the patient lies in the supine position, and the relative lengths of the inferior aspects of both medial malleoli are examined. In the supine position, the finding of a shorter leg suggests a posteriorly rotated innominate on that side; however such a finding is not confirmatory.[62] To confirm this suspicion, the examiner holds the medial malleoli with the thumbs, and the patient is instructed to come to a long-sitting position. Any apparent lengthening of the short leg confirms what was perceived as a posteriorly rotated innominate on that side.[42,62] The examiner judged that the patient's right leg appeared to be short in supine but that the legs were symmetrical in the long-sitting position. Thus the examiner judged the test to be positive on the right side, with the suspicion that the patient had a posteriorly rotated innominate on that side. The examiner next performed the prone knee flexion

test, which is another movement test that has traditionally been used to detect the side of an innominate rotation. To perform this test, the patient lies in the prone position with the shoes on, feet hanging off the plinth, and with the cervical spine at the midline. The examiner compares the difference in leg length by visually examining the difference in length of the left and right soles of the patient's shoes. A finding of one leg shorter than the other suggests that the innominate on the side of the shorter leg is in a position of posterior rotation relative to the opposite innominate; however such a finding is not confirmatory.[62] To confirm this suspicion, the examiner is supposed to passively flex the patient's knees to 90° and observe any change in the relationship of the heel positions. With passive knee flexion, an apparent increase in the leg length of the shorter leg such that it becomes equal to or longer than the longer leg is believed to indicate a posteriorly rotated innominate on that side. If this leg remains apparently shorter or becomes even shorter in relationship to the other leg, it is believed that this side is in a position of anterior rotation compared to the opposite innominate.[62] The examiner judged that the patient's left leg went from being short in the prone position with the knees extended to being longer than the right leg when the knees were flexed to 90°, suggesting a posteriorly rotated innominate on the left side according to the traditional approach. The apparent high PSIS and IC in conjunction with a low ASIS on the right in static standing suggests the presence of an anteriorly rotated innominate on the right. The results of the supine long-sitting test would suggest the presence of a posteriorly rotated innominate on the right, while the results of the prone knee flexion test suggests a posteriorly rotated innominate on the left. Clinicians who routinely use these tests to guide decision-making in the selection of an intervention such as mobilization/manipulation could reconcile the findings of the static position of the bony landmarks and the prone knee flexion test in that the presence of a posteriorly rotated innominate on one side could also be

judged as an anteriorly rotated innominate on the opposite side based on the relative position of the innominate bones.[67] However, the finding of a posteriorly rotated innominate on the right with the supine long-sitting test does not make intuitive sense in light of the results from the other tests. Some clinicians would suggest that a consistent finding of asymmetry using these tests, independent of contradictory results with respect to the side of dysfunction, may alone be sufficient to cause a clinician to consider manipulation as a potential treatment option.[62] However, given the lack of reliability and validity for these tests, it is difficult to substantiate this notion. Clinicians who frequently use these tests will attest to the notion that contradictory findings are commonplace when using these to tests to guide decision-making.

2.4 Spinal Manipulation and Low Back Pain

2.4.1 Spinal Manipulation Defined

The *Guide to Physical Therapist Practice* (*The Guide*)[68] identifies mobilization/manipulation as an intervention appropriate for the care of patients with spinal disorders. *The Guide*[68] defines mobilization/manipulation as a "manual therapy technique comprising a continuum of skilled passive movements to the joints and/or related soft tissues that are applied at varying speeds and amplitudes, including a small-amplitude/high-velocity therapeutic movement." Clinicians generally distinguish manipulation from mobilization by referring to manipulation as those techniques that involve a high-velocity low-amplitude thrust beyond the restricted range, whereas mobilization techniques are performed as relatively low-velocity, passive movements of a joint within or at the end-range of a joint's available motion.[69,70] Despite the shortcomings in the traditional approach to examine the lumbopelvic region, clinicians and researchers have

suggested that manipulation is one intervention that seems to be effective in patients with suspected dysfunction in the lumbopelvic region.[4,5,62,71-73]

2.4.2 Conflicting Evidence for Spinal Manipulation in Patients with Acute Low Back Pain

2.4.2.1 Evidence from Individual Clinical Trials

Although attempts to identify effective interventions for patients with LBP have been largely unsuccessful,[18,19] spinal manipulation is an intervention frequently used by therapists in the treatment of individuals with LBP for which there is at least some evidence to support its use. Many clinical trials have demonstrated at least some evidence for its use compared to other interventions.[1-5,7,72-102] Some studies found that patients who received spinal manipulation experienced similar outcomes related to function and disability but achieved improved satisfaction.[96,101] Although the number of studies is limited, manipulation also seems to be more effective than mobilization in the management of patients with LBP.[80] In contrast, many other trials have found little additional therapeutic benefit compared to other interventions.[6,8-11,103-116]

2.4.2.2 Evidence from Systematic Reviews of the Literature

Understandably, the conflicting findings from these trials have led to different conclusions in systematic reviews regarding the effectiveness of spinal manipulation. Some reviews suggest that spinal manipulation is helpful,[69,70,117-126] while others suggest the evidence is inconclusive because of low quality studies.[127-134] Still other reviews conclude that spinal manipulation is not helpful compared to other active interventions.[135-139]

2.4.2.3 Support in National Clinical Practice Guidelines

National clinical practice guidelines are based on summaries of evidence from clinical trials and systematic reviews. Therefore, it makes intuitive sense that enthusiasm for recommending spinal manipulation, if it is recommended at all, depends on the country.[140,141] Clinical practice guidelines in the United States,[21,142] New Zealand,[143] Denmark,[144] and Finland[145] recommend that manipulation be routinely used for patients with acute LBP (i.e. 4-6 weeks after symptom onset). Guidelines in the United Kingdom[146] and Sweden recommend it only for patients who need additional pain relief or those failing to return to normal activities.[147] Still other countries such as Switzerland[148] and Germany[149] conclude manipulation is an option, but not preferable to other treatment strategies. In contrast, guidelines in the Netherlands[150] and Australia[151] do not recommend manipulation for patients with acute LBP, and guidelines in Israel[152] concluded the evidence is inconclusive.

2.4.3 Conflicting Evidence for Spinal Manipulation in Patients with Chronic Low Back Pain

The evidence for the effectiveness of spinal manipulation for patients with chronic LBP is also unclear. Although individual RCTs[7,95] and a systematic review[126] have found some evidence for patients with chronic LBP, a recent systematic review[137] found that spinal manipulation was not effective for these patients. However, results from a recent RCT[102] not included in this review found that patients who received manual therapy demonstrated significant improvements in pain, range of motion, functional disability, general health status, and rates of return to work compared to patients who received exercise therapy. Importantly, these differences persisted at both a 6-month and one-year follow-up. Two months after treatment, 67% of patients in the manual

therapy group had returned to work versus only 27% in the exercise therapy group ($p<.01$).[102] While clinical practice guidelines in the Netherlands[150] and Denmark[144] recommend manipulation for patients with chronic LBP, most other countries do not.[21,142,143,145-149,151,152]

2.4.4 Longer-term Effectiveness for Spinal Manipulation in Patients with Low Back Pain

Unfortunately, an increasing number of clinical trials[74,76,88,94,96,100,102,106,111-113,116] and systematic reviews[126] are also beginning to report conflicting data regarding the longer-term effectiveness (i.e. at least 6 months) of spinal manipulation. Several trials[88,94,96,100,102] and one systematic review[126] demonstrate that the beneficial effects of manipulation are maintained at a longer-term follow-up period (perhaps even as long as three years[94]). However, other studies suggest that any short-term effect, if present, tends to wash out over time.[74,76,106,111-113,116]

2.5 Rationale for Conflicting Evidence for Spinal Manipulation

2.5.1 Classification of Low Back Pain

The apparently conflicting results in the clinical trials related to manipulation may be partly attributable to researchers admitting all patients with LBP into these studies, rather than selecting only those patients most likely to benefit from this intervention. Because of the inability to identify subgroups of patients with LBP based on pathoanatomical mechanisms,[23,24] attempts have been made to classify patients based on findings from the history and physical examination.[10,153-158] Identifying classification methods for patients with LBP has been recognized as an important priority both within[159] and outside the profession of Physical Therapy.[160,161] The Forum for Primary Care Research on Low Back Pain named the identification of subgroups of patients with LBP as the number one research priority in 1997.[160]

Developing effective classification rules that match patients to interventions is clearly an important priority for researchers studying patients with LBP and should improve decision-making and outcomes from physical therapy intervention by matching interventions to the patients most likely to benefit from them.[162,163] Classification methods will also enhance the power of clinical research by permitting researchers to study more homogenous groups of patients.[162,163]

2.5.2 Example of the Consequences of Ignoring Classification

The results of a systematic review regarding the effectiveness of spinal manipulation highlights the consequences of failing to adequately address the classification priority.[138] The authors conclude spinal manipulation is not effective compared to other interventions.[138] However, this definitive conclusion is a dramatic departure from their own previous reviews suggesting a benefit,[69,69,119,122,124,124,129] and they only give scant attention to whether a subgroup exists. Most disturbing, the authors attempt to immortalize the negative results by questioning the need for future clinical trials. By doing so, they ignore their own previous recommendations that future research is necessary to identify relevant subgroups.[69,124] Although the authors acknowledge a subgroup who may benefit from this intervention is "conceivable", the preponderance of the discussion suggests otherwise. This definitive position seems unwarranted when no direct effort was made in their review to consider this possibility. Rather than negating the need for future research, the results of this study beg for future studies to match individual patients to interventions with a high probability of success. Although hypothetical, if the studies in this review were more selective in their inclusion criteria, a positive effect would have more likely been detected.

2.6 Spinal Manipulation an Underutilized Intervention

Despite growing evidence for its use compared to other interventions, less than half of physicians believe manipulation is effective for patients with acute or chronic LBP.[164] Spinal manipulation is also underutilized by therapists when compared to the utilization rates of interventions that have little to no evidence to support their use.[165-167] For example, Jette and Jette[165] reported mobilization/manipulation was utilized in 35% of over 1000 patients with LBP treated by therapists. This figure does not differentiate between the use of manipulation versus mobilization, so it seems reasonable to suspect that manipulation was likely utilized in a much smaller percentage of this group of patients with LBP. On the other hand, therapists reported using other interventions with little to no evidence for their use at much higher rates, including heat/cold modalities (86%) and flexibility exercises (81%).[165] Surveys conducted among therapists outside the United States confirm the notion that manipulation seems to be underutilized among therapists worldwide.[166,167]

Li and Bombardier[166] recently surveyed 569 therapists in Canada regarding their treatment beliefs and recommendations for patients with LBP. Overall, 30% of the therapists surveyed reported that they believed spinal manipulation to be an effective treatment in the management of most patients with LBP. However, the percentage of respondents expressing a belief in the effectiveness of several non-evidence-based interventions was much higher, including ice (82%), spinal mobilization (80%), heat (66%), electrical stimulation (53%), and mechanical traction (36%).[166] In a study of 1062 patients treated for LBP in Ireland, Gracey et al[167] reported utilization of spinal manipulation in 9% of patients, compared with mobilization (44%), electrotherapy (30%), heat (19%), and traction (15%).

Unfortunately, despite widespread endorsement of spinal manipulation in several national clinical practice guidelines,[21,142-145] a recent study[168] reports that manipulation still is only used in approximately 3% of patients with LBP. The reluctance among some therapists to incorporate these skills as a routine component of their clinical practice is illustrated by the following quote: "Over the past 10 years, for example, we have seen some very compelling evidence supporting manipulation for patients with acute LBP, yet manipulation is used by therapists in typical outpatient settings at a lower-than-expected rate. What seems to be incontrovertible is the fact that evidence exists to support the use of certain treatment procedures for patients with LBP and, like other health care professionals, therapists' behavior, in many instances, does not comply with such guidelines."[169] There seem to be two primary reasons for the reluctance of some therapists to incorporate manipulation into their clinical practice.

2.6.1 Risks of Spinal Manipulation

Perhaps the number one reason therapists utilize manipulation at lower than expected rates based on the evidence for its use is related to concern regarding the potential risks.[170] In one survey of physical therapists in Canada who utilize spinal manipulation in their clinical practice, 88% of respondents strongly agreed that all available screening tests listed in the survey should be performed prior to spinal manipulation,[171] suggesting that therapists are concerned about the risks. Although little research has been conducted to quantify the risks associated with lumbar spine manipulation, manipulation of the lumbar spine is believed to be a very low risk procedure when performed by trained personnel,[172] and the risk of a serious adverse event appears to be rare.[125] Although the precise occurrence rate of serious complications of manipulation to the low back is unknown, the most serious complication reported in the literature is cauda equina

syndrome, which has only been reported in a very few number of cases.[173] In a review of the literature from 1911-1989, Haldeman et al[173] found ten reported cases of cauda equina syndrome after manipulation of the lumbar spine over a 77-year period of published literature. In an attempt to improve the precision of this estimate, investigators[124,174] have estimated the incidence of cauda equina syndrome to be on the order of less that one per 100 million manipulations. It is believed that the risk of cauda equina syndrome may be the greatest when manipulation is performed in the presence of sciatica or under the influence of anesthesia.[124]

2.6.1.1 Less Serious Side Effects Associated with Spinal Manipulation

Senstad et al[175] studied the characteristics of less serious but "unpleasant" side effects of 4712 manipulative treatments across all body regions in 1058 patients treated by chiropractors in Norway. Importantly, there were no adverse events reported as serious or life-threatening in this study, but 55% of the patients interviewed reported at least one side effect from treatment. The most common side effects included local discomfort (53%), local headache (12%), fatigue (11%), or radiating discomfort (10%). Patients characterized 85% of these complaints as "mild" or "moderate", with 64% of side effects appearing within four hours after manipulation. Within 24 hours after manipulation, 74% of the complaints had resolved. Less than 5% of side effects were characterized by dizziness, nausea, hot skin, or "other" complaints. Side effects were rarely still noted on the day after manipulation, and very few patients reported the side effects as being severe. Perhaps most importantly, patients rarely reported that the side effects resulted in a reduction in their activities of daily living. Leboeuf-Yde et al[176] reported on 1858 spinal manipulations performed on 625 patients by chiropractors in Sweden. In this study, 44% of patients reported experiencing local discomfort, headache or fatigue at least once, and in 19% of these cases these symptoms persisted for greater than 48 hours. No serious complications were

reported.[176] Neither of these studies distinguished manipulations performed on the lumbar spine from those performed on other regions of the spine.

2.6.1.2 Risks of Spinal Manipulation Compared to Non-steroidal Anti-inflammatory Medication

To place the risk of lumbar spine manipulation in perspective, it is useful to consider the potential side effects of one of the most common treatments for acute LBP, non-steroidal anti-inflammatory medication (NSAID). Thirty percent of patients experience at least one side effect from NSAID use, with this risk further increasing with prolonged use, defined as greater than four weeks.[177] One to three percent of patients who take a NSAID are believed to experience gastrointestinal (GI) bleeding secondary to NSAID use.[178] The risk of serious GI complications from taking a NSAID is 3.2 per thousand in individuals over the age of 65 and one per thousand individuals when considering all ages.[179] Perhaps most alarming, Tamblyn et al[180] reports that approximately 7,600 deaths and 76,000 hospitalizations may be attributable to NSAID use each year in the United States. Based on the nature of the risk of NSAID use, the risk of manipulation appears to be quite low and well within reason given its demonstrated efficacy.

2.6.2 Spinal Manipulation an "Advanced" Skill?

The second reason some clinicians may be reluctant to consider spinal manipulation for patients with LBP is based on the impression that it is an advanced skill that requires a high level of skill and practice. Empirically, a therapist's skill level is believed to be related to clinical outcome. A recent survey of licensed physical therapists demonstrated that, on average, therapists perceive spinal manipulation as an advanced skill to be acquired through post-professional, as opposed to entry-level education.[181] Therapists also perceive that incorrect performance of these

interventions will cause "severe psychological or physical harm".[181] Similar data exists regarding the views toward manipulation of the extremity joints.[181] However, there is little data to support these somewhat alarmist views. Studies[182-184] have shown that with practice of a task, newly trained practitioners are able to apply similar levels of force compared to skilled practitioners. Increased practice has also been shown to improve performance regardless of experience,[184] reinforcing the notion that spinal manipulation is a motor skill that simply requires repetition. Future work from this study will specifically examine the relationship between experience and treatment outcome.

Additionally, many theoretical approaches to identify patients likely to benefit from spinal manipulation have been proposed;[156,185-191] however, there is little to no evidence to support any single approach. These approaches frequently incorporate complex diagnostic schemes based on pathoanatomical and biomechanical theories that utilize various examination procedures to identify a pathological motion segment, or a biomechanical dysfunction towards which a manipulative intervention is then directed. However, research has shown that relevant pathoanatomical mechanisms can only be identified in a small percentage of patients with LBP,[23,24] and many of the tests proposed to identify biomechanical dysfunction are of questionable reliability,[32-34] and validity.[27-30] negating any face validity they seem to possess. In fact, it has been suggested that clinicians entirely abandon the notion of detecting spinal position in light of the overwhelming evidence that these methods are largely unreliable and not valid.[36,37] Many of the more commonly used diagnostic schemes seem to require a large degree of "mental gymnastics" but offer little benefit to the patient with LBP who would otherwise benefit from a clinician with the attitude to "move it and move on".[192] Recent evidence[193] also questions the

value of specific motion testing for decision-making regarding the use of manipulation in patients with neck pain.

Continuing education courses that teach these approaches may incorporate a large number of techniques and perhaps only complicate the situation. Clinicians may be falsely led to believe they need to become familiar with many techniques before they can be considered proficient using any single one. They may be left with the notion that there are an infinite number of biomechanical patterns, each of which suggests a unique manipulative intervention should be used. Although a complex diagnostic and treatment classification scheme may validate our need to demonstrate a high level of intellectual and diagnostic prowess, there is little evidence to suggest that any of these schemes lead to greater improvements in outcome than the use of more general manipulative interventions that can be used in a wide spectrum of patients by the entry-level therapist. This is not to suggest that decision-making in the use of spinal manipulation should be haphazard or random, or that therapists can be cavalier in their approach; however, examination findings and models of decision-making must be considered only in the context for which there is evidence to support their use. Advanced training through professional continuing education and residency/fellowship training clearly has some value for clinicians who desire to become increasingly proficient with a variety of manual interventions. Such training may improve a clinician's diagnostic, decision-making, and intervention skills to manage a wider and more complicated spectrum of patients with spinal dysfunction. However, the unintended consequence of a dogmatic, non evidence-based approach that incorporates an overwhelming number of manipulative interventions is that rather than encouraging the use of spinal manipulation, many clinicians may never apply these skills in their clinical practice. Results

from clinical trials[4,5] and findings from a study to identify the characteristics of patients likely to benefit from a spinal manipulation[12] suggests that many patients with LBP may benefit from a single manipulative intervention.

2.7 An Alternative Approach in the Management of LBP

Based on the important research priority of developing meaningful classification systems for patients with LBP, identifying the relevant subgroup of patients with LBP likely to benefit from spinal manipulation is one step that can help to minimize such conflicting results from occurring in research that assesses the efficacy of manipulation in patients with LBP. It appears that an accurate identification of patients with LBP with suspected dysfunction in the lumbopelvic region is essential in selecting a proper and effective treatment, thus many clinicians who treat LBP regularly assess this region for dysfunction. However, the lack of evidence for the current approach to identify this subgroup of patients may in fact make it impossible to even identify who these patients are using the pathoanatomical approach. However, clinicians are still confronted with the evidence that manipulation seems to be an effective intervention in some patients with LBP. The futility of the traditional approach in combination with evidence that suggests that spinal manipulation is effective requires clinicians and researchers seek an alternative methodologic paradigm to identify this subgroup of patients.

2.7.1 Limitations of the Randomized Clinical Trial

The RCT is the highest level of evidence to elucidate an intervention's effect.[163] Based on many rigorous characteristics of the RCT that are beyond the scope of this discussion, the RCT is subject to the least amount of bias, thus clinicians can have a high degree of confidence that the demonstrated treatment effect may be attributed to the intervention rather than some other known

or unknown factor. For example, several national clinical practice guidelines[21,142-145] recommend spinal manipulation for patients with non-radicular acute LBP based on reviews of RCT evidence for this intervention..

Despite the high level of evidence for RCTs, they are conducted on groups of patients who are randomly assigned to a treatment arm, thus the results are ideally suited to assist decision-making for groups of patients.[194] In other words, when examining large groups of patients undergoing treatment for LBP, spinal manipulation should be observed to be in frequent use by clinicians who use evidence from the literature to guide decision-making. However, clinicians obviously do not treat groups of patients, thus clinical practice guidelines are ineffective in helping clinicians determine if the individual patient sitting in front of them might benefit from a this intervention.

For example, suppose a clinician wants to conduct a RCT comparing the effect of spinal manipulation and exercise versus exercise alone to improve pain and function in patients with LBP. To examine the hypothesis that patients who receive spinal manipulation will experience greater improvements than patients who receive only exercise, the researcher might design a RCT in which patients are randomly assigned to one of these two groups. Classic inferential statistical procedures such as t-tests and the analysis of variance would be used to compare the groups on the outcomes of interest. Suppose the results demonstrate that patients who receive spinal manipulation improve an average of 30 points after four weeks, versus a 10-point improvement for patients who receive only the stabilization exercise intervention. Because the results can only be generalized to how groups of patients will respond on average and not useful

to estimate an individual patient's prognosis, clinicians are unable to counsel their next patient that he or she can expect to improve 30 points if the same intervention is provided. This limitation of the RCT makes intuitive sense when the methodology behind classic hypothesis testing is examined. In its most simple form, the means of the outcome variables between the groups are compared. Means summarize the average effect of the intervention, thus do not describe how an individual response contributes to the overall mean, unless of course every patient in the sample responded in a similar fashion. In other words, unless patients are similar to the average patient in the sample, there is no way to specifically counsel them as to their chance of improvement given exposure to the intervention being studied. The mean response has even less value in situations where the variability in responses to a particular intervention is quite disparate. Therefore, when counseling patients on the effectiveness of interventions based on RCT evidence, clinicians can only use phrases such as, "on average." This is not to suggest that RCTs are not useful. In fact, quite the opposite is true. However, the RCT is not the final answer as to whether a particular intervention may be beneficial for an individual patient sitting in the waiting room. It seems reasonable to suspect that the inability to identify individual patients who might benefit from spinal manipulation has contributed to its persistent underutilization,[165-167,170] despite widespread endorsement for its use.[21,21,142,143,143-146]

2.7.2 Clinical Prediction Rules

2.7.2.1 Examples of Clinical Prediction Rules in the Literature

Clinical prediction rules (CPR) are tools designed to assist in the classification process and improve decision-making for clinicians caring for individual patients.[195,196] Historically, CPRs have been developed to improve the accuracy in making a diagnosis and establishing a patient's

prognosis.[195,196] For example, for musculoskeletal disorders, CPRs have been developed to improve the accuracy of diagnosing ankle fractures (i.e. "the Ottawa ankle rules")[197] and knee fractures (i.e. "the Ottawa knee rules")[198] in individuals with acute injuries. CPRs have also been developed to determine when to order CT in patients who have experienced minor head injuries[199,200] and radiographs in patients who have experienced neck trauma.[201] CPRs have also been developed to more accurately diagnose strep throat,[202] coronary artery disease,[203] and pulmonary embolism,[204] for example. To establish prognosis, researchers have developed CPRs to determine when to discontinue resuscitative efforts after cardiac arrest in the hospital,[205] determine the likelihood of death within four years for individuals with coronary artery disease,[205] identify children at risk for developing urinary tract infections,[206] and identify the characteristics of patients likely to develop post-operative nausea and vomiting after anesthesia.[207] A CPR has recently been developed to establish the prognosis of patients who have experienced a rear-end motor vehicle accident.[208] Although CPRs can be developed to improve the accuracy of making a certain diagnosis or establish patient prognosis, this discussion will focus on the development of a CPR to predict patients likely to benefit from a specific intervention based on the outcome from treatment. However, the methodology is similar for both purposes.

2.7.2.2 The First Step: Creating the Clinical Prediction Rule

The first step in the development of a CPR involves creating the rule. This requires the researcher to examine the ability of multiple factors from the history and physical examination to predict an outcome of interest. The outcome of interest serves as the reference criterion or "gold standard" by which a successful outcome will be judged. The selection of an appropriate and clinically meaningful reference criterion is pivotal to the usefulness of the CPR that is eventually

developed. All possible factors that are believed to be related to the outcome of interest should be included as potential predictors. These predictors may be selected based in the researcher's clinical experience and/or previous research that have demonstrated a particular factor or set of factors to be related to the outcome of interest. While it may seem beneficial to simply include every possible factor from the history and physical examination, the researcher must weigh the benefits of including a complete set of potential predictor variables against the increase in sample size required for each additional variable that is added as a potential predictor. Once the set of predictor variables is established, subjects are exposed to the treatment of interest and then judged to be either a success or non-success against the reference criterion based on a pre-determined clinically relevant cut-off score. Although other techniques may be used, logistic regression is a commonly used statistical approach that can then be used to determine the most powerful set of predictor variables to maximize accuracy.[195] The details of how to conduct logistic regression are beyond the scope of this discussion, but the reader is referred to Kleinbaum et al[209] for a more detailed discussion of this topic. Classic hypothesis testing involves the comparison of group means using traditional statistical procedures such as the analysis of variance. In contrast, the development of CPRs utilizes diagnostic properties of sensitivity, specificity, and positive and negative likelihood ratios, which are based on the individual patient. Thus their interpretation can be readily applied to an individual.

2.7.3 The Need for Clinical Prediction Rules in the Management of Low Back Pain

A CPR could also be developed to improve the accuracy of the classification of patients with LBP. The classification process is used to identify patients with particular characteristics who will likely benefit from a specific type of treatment. Clinicians who frequently use spinal manipulation will attest to the notion that some patients experience relatively dramatic

improvements after only one to two treatments. Given the precarious evidence for the traditional approach to identify patients with dysfunction in the lumbopelvic region, it seems that a more successful approach might be to develop a CPR to identify patients likely to achieve a relatively dramatic improvement in only a short period of time using an intervention like spinal manipulation, which at least some evidence for its use. A few RCTs have found manipulation to be more beneficial for a subgroup of patients with more acute symptoms,[80,84] or more limited straight leg raise range of motion.[83] In the previous studies that suggested manipulation to be an effective intervention for some patients with LBP,[4,5] the criteria for classifying patients as manipulation candidates was based strictly on clinical experience and relied heavily on the results of diagnostic tests for dysfunction in the lumbopelvic region, which, as previously discussed, are not useful for decision-making because of their inherently poor reliability and validity. However, none of these studies examined the addition of other important variables from the history and physical examination that would maximize the prediction of success with manipulation prior to the intervention. In a systematic review of the literature, Di Fabio[123] suggested a preliminary "profile" of common characteristics among patients who seem to benefit from spinal manipulation, which consisted of the following factors: 1) acute symptom onset with less than a 1-month duration of symptoms, 2) central or paravertebral pain distribution, 3) no previous exposure to spinal manipulation, and 4) no pending litigation or workers' compensation. Although spinal manipulation appears to be effective for some patients, little systematic efforts have been undertaken to identify the characteristics of patients with LBP likely to likely benefit from this intervention.

2.7.4 The Development of a Spinal Manipulation Clinical Prediction Rule

A CPR to accurately predict which patients will most likely benefit from spinal manipulation would be immensely helpful for decision-making. Despite the prevalence of LBP and the inherent difficulties in selecting effective interventions, little work has been done to establish such a rule. Flynn et al[12] attempted to determine criteria from the history and physical examination in patients with LBP that would predict patients likely to likely benefit from spinal manipulation. The results of this study, which were recently published in *Spine*,[12] accomplished the first step in the development of a CPR by creating the CPR. Details of how the study was conducted are outlined below.

2.7.4.1 Inclusion/Exclusion Criteria

All patients between 18 and 60 years of age with non-radicular LBP who agreed to participate in the study received a standardized examination of their spine. Patients had to have at least 30% disability on the modified ODQ[210] to be admitted.

Patients who were pregnant, exhibited signs consistent with nerve root compression (i.e. positive straight leg raise at less than 45^0, diminished lower extremity strength, sensation, or reflexes, etc.), history of prior lumbar spine surgery, or a history of osteoporosis or spinal fracture were excluded from the study.

2.7.4.2 Self-report Measures

Patients admitted to the study completed a series of self-report measures during a baseline examination that included demographic information, an 11-point numeric pain rating scale,[211]

and a pain diagram[212] to determine the most distal extent of their symptoms.[213] Patients also completed the ODQ and the Fear-Avoidance Beliefs Questionnaire (FABQ).[214] The ODQ is a self-report measure of function commonly used in patients with LBP.[215-219] The questionnaire consists of ten items addressing different aspects of function, each scored from 0-5 with higher values representing greater disability. The ODQ used in this study was the modified version, which replaced the section on sex life with one regarding employment/home-making. Previous research has found the modified version to have high levels of reliability, validity and responsiveness.[210] The FABQ is made up of two subscales and is used to assess the extent to which patients believe physical activity (Physical Activity Subscale) and work (Work Subscale) influences their LBP.

2.7.4.3 History and Physical Examination

After completing the questionnaires, the examiner conducted a standardized history and physical examination. Patients were asked about the mechanism of injury, nature of current symptoms, and prior episodes of LBP. A neurologic screening examination was conducted to rule out any evidence of nerve root compression or radiculopathy, which is generally viewed to be a contraindication for manipulation,[21] and was an exclusion criteria in this study. Testing included sensory testing, motor testing, muscle stretch reflex testing, and tension signs such as the straight leg raise and prone knee flexion tests. The examination also included Waddell's nonorganic signs.[220] Range of motion and status change in symptoms with single, cardinal-plane lumbar movements was recorded.[156] Supine straight leg raise and hip internal and external range of motion were measured. Segmental mobility testing of the lumbar spine was conducted to assess pain provocation and segmental mobility.[186,221] Each segment was judged to be normal, hypomobile, or hypermobile. Importantly, many of the traditional diagnostic tests purported to

identify dysfunction in the lumbopelvic region were also performed. These included tests designed to assess the symmetry of bony landmarks in the static position (static symmetry tests), tests to assess the symmetry of bony landmarks with movement (movement symmetry tests), and tests to reproduce symptoms (provocation tests). The operational definition used for each of these tests and the criteria to judge a positive test are included in Appendix A.

2.7.4.4 Manipulative Intervention

Once the examination was complete, all patients received a general manipulative intervention purported to affect the lumbopelvic region. Regardless of the clinical examination findings, all patients received the same intervention. To perform the manipulation, the patient is placed in the supine position, and the clinician stands on the side opposite of that to be manipulated. The following decision rule was used to determine the side to be manipulated: 1) the side of the positive standing flexion test; 2) if this test was negative, the side of tenderness during sacral sulcus palpation was selected; 3) if neither side was tender, the more symptomatic side base on the patient's self-report was selected; and 4) if none of these criteria were satisfied, the clinician flipped a coin to determine the side to be manipulated.[12] Although the manipulative intervention is directed towards one side of the pelvis, Cibulka et al[222] found changes in innominate tilt on both sides of the pelvis after the performance of this manipulation on one side. Therefore, it is likely that the manipulation impacts both sides of the lumbopelvic region. The patient is passively moved into side-bending towards the side to be manipulated. The patient interlocks the fingers behind his or her head. The clinician passively rotates the patient, and then delivers a quick thrust to the anterior superior iliac spine in a posterior and inferior direction. The clinician noted whether a cavitation was heard and felt. Based on the presence of a cavitation, no further manipulation was provided. If no cavitation was produced, the procedure was to reposition the

patient and then attempt the manipulation again.[12] If no cavitation was experienced, the clinician attempted to manipulate the opposite side. A maximum of two attempts per side was permitted. If no cavitation was produced after the fourth attempt, the clinician proceeded to instruct the patient in a range of motion exercise. A video clip of the manipulative intervention can be viewed by clicking on video clip (Figure 1).

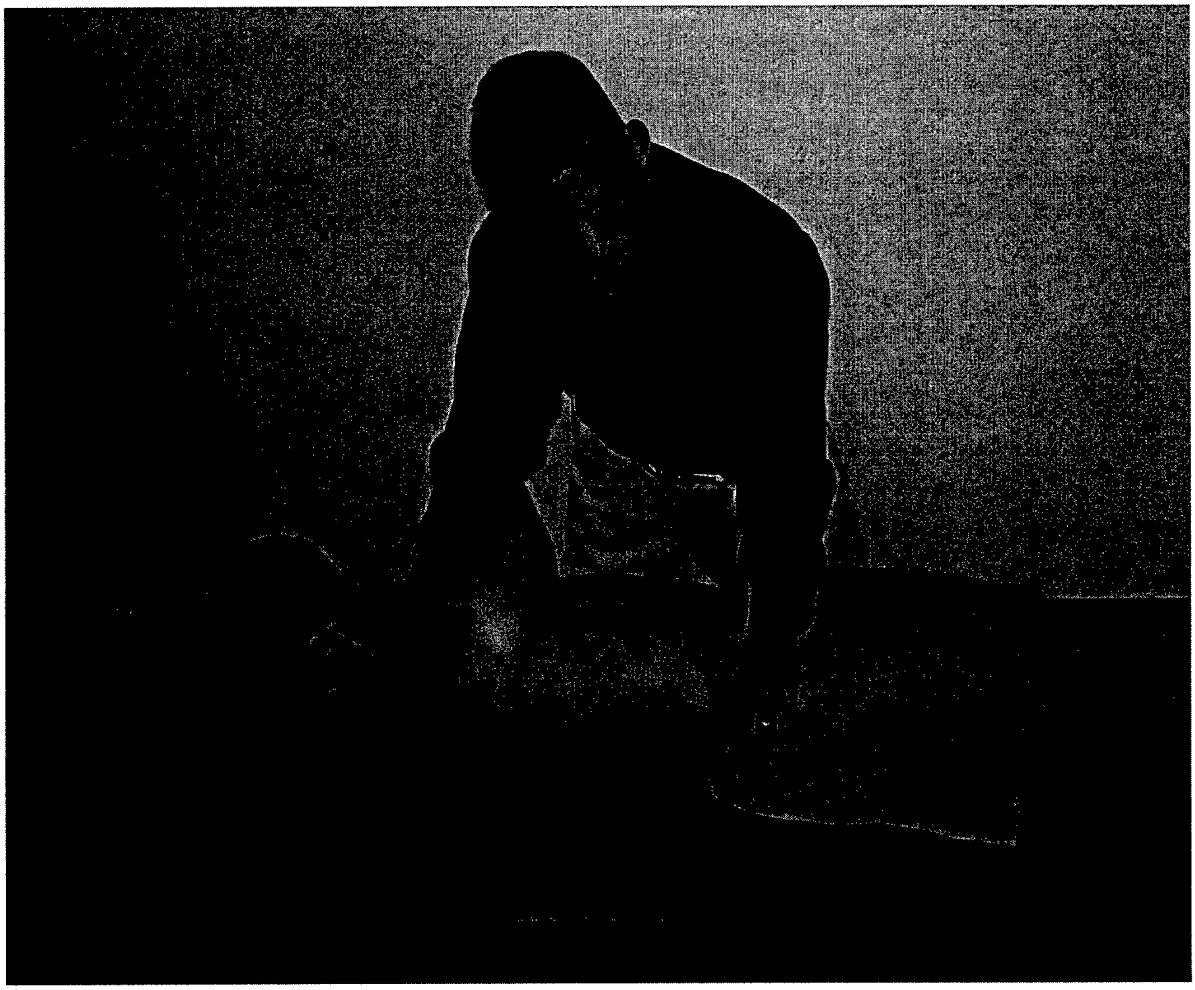


Figure 1. Specific manipulative intervention used in the development of the spinal manipulation CPR.

(Return to p. 87)

Following the manipulative intervention, the patient was instructed in a supine pelvic tilt exercise. The patient performed a set of 10 repetitions in the clinic and was instructed to perform 10 repetitions of the exercise 3-4 times daily. Finally, the patient was instructed to maintain usual activity level within the limits of pain. Advice to maintain usual activity has been found to assist in recovery from LBP.[21,223] The patient was instructed to do all activities that did not increase symptoms and to avoid activities which aggravated symptoms.

2.7.4.5 Determination of Successful Outcome

The next step was to classify patients according to the reference criterion. Reid et al[224] suggest that an appropriate reference criterion is one that represents the condition in which the diagnostic test is attempting to identify. The purpose of this study[12] was to "diagnose" patients with LBP likely to benefit from spinal manipulation using relevant factors from the history and physical examination. Therefore, outcome from manipulation was used as the reference criterion to identify patients who succeeded. Using a clinically relevant reference criterion such as treatment outcome is a dramatic improvement over the limitations imposed by previous studies that used short-term pain relief after lumbopelvic region injection[25,44,60,61,63-65] and the presence or absence of LBP[62] as the reference criterion.

Previous research has shown an average ODQ score of approximately 40% for new patients referred to physical therapy, with a standard deviation of about 10%.[216,225] Because change in disability was used as the reference standard,[12] patients were required to have a minimum baseline score of 30% on the ODQ. This minimum baseline level of disability insured the inclusion of a spectrum of patients, yet prevented a floor effect from occurring due to low baseline disability scores. Additionally, the minimum clinically important difference (MCID) for

the ODQ has been shown to be 6 points.[210] A threshold of 30% on the ODQ at baseline using 50% improvement as the reference standard to judge success helps clinicians to distinguish patients who are responding to the manipulative intervention versus improvements solely attributable to the favorable natural history of LBP.[219] At a minimum threshold of 30%, a 50% improvement corresponds to a 15-point improvement in disability. This magnitude of change represents 2.5 times the MCID for the ODQ, increasing the clinician's confidence that the improvement can be attributed to the manipulative intervention rather than a favorable natural history. Higher baseline levels of disability result in even greater magnitudes of improvement when using 50% improvement as the reference standard.

Based on the response to spinal manipulation as the reference criterion, patients were judged at the second visit 2-4 days later as to whether they experienced a successful outcome. The purpose of the CPR is to identify patients who experience important clinical changes in disability in a relatively short period of time, changes beyond that likely attributable to the favorable natural history of LBP. These are the patients clinicians surely do not want to miss when considering the use of spinal manipulation. Because of the desire to protect against the argument that improvement in outcome could merely be associated with the favorable natural history of LBP, a relatively high threshold was established to differentiate those patients judged to be a success and those patients judged to be a non-success. Previous research[4,5,225] that utilized the same manipulative intervention has demonstrated that a 50% improvement in the ODQ is able to distinguish between patients responding to manipulation versus those simply benefiting from the favorable natural history of LBP. This reference criterion seems even more rigorous given that improvement in this previous research occurred across a 1-4 week period, whereas the

improvement in the study that developed the CPR was expected to occur in only a few days. Clinicians who routinely use the ODQ will also admit that a 50% reduction in the ODQ in this short period of time is a relatively dramatic improvement.

Patients who improved at least 50% were then out of the study and judged to be a success. Those who did not achieve at least a 50% improvement were manipulated again and then re-checked at a third visit 2-4 days after the second visit. Patients were once again judged as to whether the manipulation was a success, based on the reference criterion of a 50% improvement in ODQ from the initial visit. Patients who did not achieve at least a 50% improvement in ODQ at this point were categorized as a non-success.

2.7.4.6 Data Analysis

A total of seventy-one patients completed the study. All of the variables included in the self-report questionnaires and the history and physical examination were initially assessed in a univariate fashion with an α-level equal to 0.15 to determine their individual accuracy in predicting success. Multivariate logistic regression was then performed on all the variables demonstrated to be significant at the 0.15 level in a univariate fashion to determine the most parsimonious set of factors from the history and physical examination to identify patients who achieved at least a 50% improvement in the ODQ.

2.7.4.7 Criteria in the Spinal Manipulation Clinical Prediction Rule

The multivariate analysis resulted in a CPR that consisted of five criteria, each of which can be easily assessed in the clinic. The criteria and threshold for determining whether a patient is positive with respect to each criterion is listed in Table 1.

Table 1. Criteria in the Spinal Manipulation CPR.

Criterion	Definition of positive
1. Duration of current episode of LBP	< 16 days
2. Extent of distal symptoms	Not having symptoms distal to the knee
3. FABQW subscale score	< 19 points
4. Segmental mobility testing	At least one hypomobile segment in the lumbar spine
5. Hip internal rotation range of motion	At least one hip with > 35° of internal rotation range of motion

(Return to p. 68)

The traditional diagnostic tests used to assess for dysfunction in the lumbopelvic region were uniformly either not sufficiently reliable and/or not related to outcome, which should cause clinicians to further question the utility of these tests for decision-making. Appendix B includes the data with respect to the percentage of diagnostic tests that were positive in each group, and the univariate significance level for each test. In essence, any clinical factor that could possibly be related to outcome from spinal manipulation was included in the examination to minimize the chance of excluding a potentially important variable that may be prognostic of the response to manipulation.

2.7.4.8 Accuracy of the Spinal Manipulation Clinical Prediction Rule

The accuracy of the CPR can be expressed using likelihood ratio (LR) statistics. The positive LR expresses the change in odds favoring the outcome when the patient meets the criteria in the CPR, while the negative LR expresses the change in odds favoring the outcome when the patient does not meet the rule's criteria.[226] An accurate CPR should therefore have a large positive LR or a small negative LR. According to Jaeschke et al[227] accuracy can be considered moderate when

the positive LR is greater than 5.0 or the negative LR is less than .20. Accuracy is substantial when the positive LR is greater than 10.0 or the negative LR is less than .10.[227] Because this study sought to identify patients who would likely benefit from spinal manipulation, the statistic of interest was the positive LR.

Forty-five percent (32/71) of patients were classified as a success, regardless of patient's status with respect to the CPR.[12] In other words, if clinicians were to randomly manipulate patients with non-radicular LBP, they can expect to achieve at least a 50% improvement in the ODQ by the end of approximately one week 45% of the time. However, when considering a patient's status with respect to the CPR, the positive LR was 24.4 for patients who met at least 4/5 criteria. To put this result in perspective, the probability of achieving a successful outcome increases from 45% to 95%. With three factors, the positive LR was 2.6, which translates into a 68% probability of success. Given the ease with which the CPR is applied and manipulative intervention can be performed, and in light in the extremely low risks,[124,173,174,176] this is still likely a sufficient probability to justify an attempt at manipulation. With less than three criteria met, the probability of success is essentially no better than the probability of success if you were to randomly manipulate patients with non-radicular LBP. Thus the clinician may want to consider other interventions that have a higher probability of success.

2.7.4.9 A Traditional Versus an Evidence-based Approach for Decision-making

Two case reports were recently published that outline how to use the CPR to improve decision-making to identify patients likely to benefit from spinal manipulation compared to the traditional approach using the classic diagnostic tests to assess dysfunction in the lumbopelvic region.[66] The examination of each patient included a standardized history and physical examination similar to

those described in the initial study that developed the CPR[12] to facilitate the determination of each patient's status with respect to the CPR.

2.7.4.9.1 Case Description – Patient #1

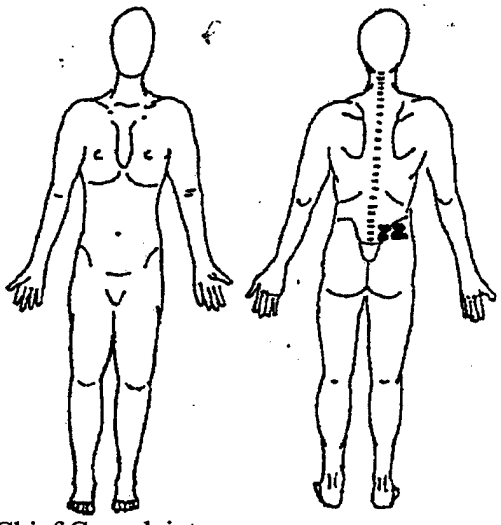
2.7.4.9.1.1 History and Self-Report Measures

The first patient was a 54 year-old male with a history of more than 10 episodes of LBP over the previous five years (Table 2).

Table 2. Key findings from the self-report measures and history.

(Return to p. 59, 61, 61)

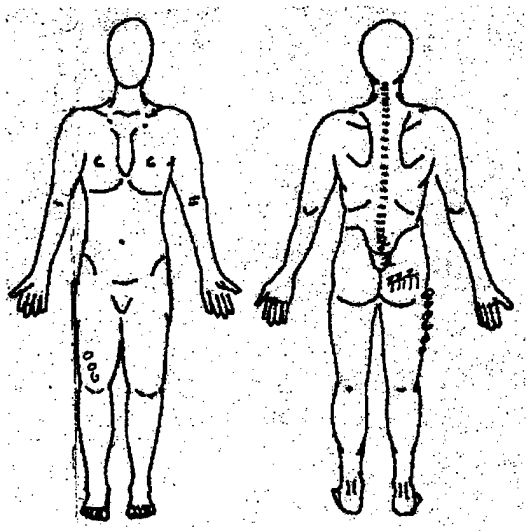
Patient #1



Chief Complaint:

Right-sided low back pain	
Duration of Symptoms:	Symptoms began gradually 1 week ago
Prior History of Back Pain:	Approximately 10 prior episodes of low back pain over the past 5 years
Pain Rating:	Current level of pain 5/10
Oswestry Score:	42%
FABQ – Work Subscale Score:	8/42

Patient #2



Chief Complaint:	Right buttock pain, pain and numbness in right anterior/ lateral thigh
Duration of Symptoms:	Symptoms began 3 years prior while running track in college
Prior History of Back Pain:	No previous history of low back pain
Pain Rating:	Current level of pain 4/10
Oswestry Score:	26%
FABQ – Work Subscale Score:	Not assessed

Previous treatment included a heel lift, spinal manipulation, and lumbar spine stabilization exercises, to which he had previously responded positively. Based on his frequent history of LBP and positive response to lumbar stabilization exercises, he was suspected to have segmental instability.[228] The most recent episode started gradually one week ago and had worsened over the last 2-3 days. He did not recall a specific mechanism of injury. His predominant symptom was right-sided LBP in the area of the lumbopelvic junction. The pain diagram and results of the self-report measures are provided in Table 2. The ODQ score (42%) revealed moderate disability and the FABQ-work subscale score (8/42) showed a low level of fear-avoidance beliefs.

2.7.4.9.1.2 Physical Examination

The results of the physical examination are summarized in Table 3.

Table 3. Key physical examination findings.

Examination Component	Patient #1	Patient #2
1. Neurologic Screening	No strength, sensory or reflex changes	No strength, sensory or reflex changes
2. Lumbar Active range of motion and Status Change		
Flexion	85^0, status quo	78^0, status quo
Extension	30^0, status quo	22^0, status quo
Right Side-Bending	25^0, status quo	25^0, status quo
Left Side-Bending	15^0, status quo	31^0, status quo
3. Hip Rotation Passive range of motion		
Right Hip Internal Rotation	45^0	18^0
Right Hip External Rotation	45^0	35^0
Left Hip Internal Rotation	35^0	25^0
Left Hip External Rotation	40^0	32^0
4. Lumbar Segmental Mobility and Pain Provocation	Hypomobility and pain provocation at L_4 and L_5	Mobility judged to be normal at all levels, pain provocation at L_5
5. SI Symmetry Tests		
PSIS symmetry standing	Right side judged to be higher	Left side judged to be higher
ASIS symmetry standing	Left side judged to be higher	Right side judged to be higher
Iliac crest symmetry standing	Right side judged to be higher	Judged to be symmetrical
PSIS symmetry sitting	Judged to be symmetrical	Right side judged to be higher
Iliac crest symmetry sitting	Judged to be symmetrical	Judged to be symmetrical
6. SI Mobility Tests		
Standing Flexion Test	Positive on the right	Positive on the right
Seated Flexion Test	Positive on the right	Positive on the right
Gillet Test	Positive on the right	Positive on the right
Supine Long-Sitting Test	Positive	Negative
Prone Knee Flexion Test	Positive	Not Assessed

(Return to p. 61, 62)

The results of the neurologic examination did not reveal any sensory, strength, or reflex deficits. There was no centralization or peripheralization noted during lumbar range of motion testing. Rotation range of motion of the left hip was somewhat limited compared to the right hip. During segmental mobility testing, the L_4 and L_5 segments were judged to be hypomobile and pain was also provoked. There were numerous positive findings for dysfunction in the lumbopelvic region, including the standing and seated flexion tests, the Gillet, supine long-sitting, and prone knee flexion tests (Table 3).

2.7.4.9.2 Case Description – Patient #2

2.7.4.9.2.1 History and Self-Report Measures

The second patient was a 26 year-old male, with complaints of right-sided buttock pain and intermittent pain and numbness into the right anterior/lateral thigh (Table 2). These symptoms had begun approximately three years prior to the examination while the patient was a sprinter for his college track team. The onset was gradual, and the symptoms prevented him from running at the time of the examination. The ODQ score (26%) revealed a lower level of disability for this patient (Table 2). A baseline score of 26% does not strictly meet the minimum 30% level of disability used in the initial study that developed the CPR;[12] however, a score of 26% is only .5 standard deviations below this minimum.[216,225] Additionally, based on our clinical experience, 26% on the ODQ is still a sufficient level of disability to prevent a floor effect from occurring, despite falling below the 30% threshold. Most importantly, a 50% reduction in the ODQ score for a patient with a baseline score of 26% still represents a clinically important improvement in disability.[210] The FABQ was not assessed on this patient.

2.7.4.9.2.2 Physical Examination

The results of the physical examination are summarized in Table 3. Similar to the first patient, the results of the neurologic examination were negative and there was no peripheralization or centralization noted during lumbar range of motion. Hip range of motion was generally less than patient #1, and the right hip appeared to be limited in internal rotation as compared to the left hip. The right hip was also limited in flexion range of motion as compared to the left hip. Segmental mobility testing provoked pain at the L5 segment, and the mobility was judged to be normal at all lumbar levels. There were also several positive findings for dysfunction in the lumbopelvic region including the standing and seated flexion tests, as well as the Gillet test.

2.7.4.9.3 Clinical Decision-Making Based on Traditional Diagnostic Tests

Both patients had several positive findings on traditional tests designed to detect dysfunction in the lumbopelvic region, including the standing and seated flexion tests, and the Gillet test. Patient #1 also had positive findings on the supine long sitting and prone knee flexion tests. Both patients were judged to have asymmetry of the pelvic landmarks, which is often believed to indicate dysfunction in the lumbopelvic region.[4,5,62] Patient #2 had signs of hip joint dysfunction as well as possible dysfunction in the lumbopelvic region. Based on these results, both patients appeared to be good candidates for spinal manipulation directed at the lumbopelvic region, and were treated in this manner at the first appointment.

2.7.4.9.4 Clinical Decision-Making Based on Clinical Prediction Rule

Each patient's status with respect to the CPR[12] is outlined in Table 4.

Table 4. Status of the two patients with respect to the CPR.

Criteria in the CPR	Patient #1	Patient #2

1. FABQW score < 19 points	8	Not assessed
2. Duration of current episode < 16 days	5 days	3 years
3. No symptoms extending distal to the knee	Low back pain only	Right buttock/ thigh pain, not distal to the knee
4. At least one hypomobile lumbar spine segment (judged from lumbar spring testing)	Hypomobility at L_4 and L_5	Mobility judged WNL at all lumbar levels
5. At least one hip with > 35° of internal rotation range of motion	Left Hip IR - 35° Right Hip IR - 45°	Left Hip IR - 25° Right Hip IR - 18°
TOTAL	**5/5**	**1/5**

(Return to p. 68, 68)

The first patient met all five criteria in the CPR, which suggests that he is highly likely to achieve a dramatic improvement with the manipulative intervention. The second patient met one criterion (no symptoms distal to the knee). This patient was not assessed on the FABQ at baseline, therefore his score of the work subscale could not be factored into the CPR. Patient #2 may therefore have met a maximum of two criteria. In either case, patient #2 met two or fewer criteria, making him unlikely to experience dramatic improvement with this manipulative intervention.

2.7.4.9.5 Interventions and Outcomes

Based on the results of the traditional lumbopelvic region tests, both patients were treated using the manipulative intervention that has been previously described in the initial study that developed the CPR,[12] and in earlier studies that have shown this particular technique to be effective.[4,5] Because of the lack of reliability in the judgments from tests often used to determine

which side to manipulate,[52] the more symptomatic side was selected. An audible cavitation was achieved in both patients using the technique. Following the manipulation, each patient was instructed in a hand-heel-rock range of motion exercise as described in Figure 2.[229]

Performance:
Starting Position
Get on all fours on the floor. Rest some of the weight on your hands and arms; move your hands to just slightly higher than your shoulders.

Forward Rock
Transfer the weight more to your hands, not allowing your arms to bend. Allow your abdomen to sag towards the surface while your head tends to look up. Pause momentarily toward the end of your range and then start back towards neutral.
Backward Rock
Rock backwards as though you were attempting to sit on your heels. Allow your back to round out and do not be concerned if you have to drag your hands along the surface to get back to the fully backward position.

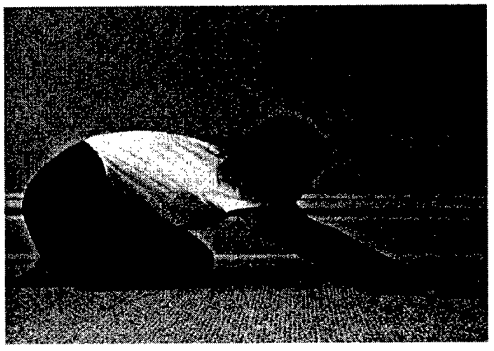


Figure 2. Description of the hand-heel rock range of motion exercise.

We routinely instruct patients in this exercise to help maintain immediate improvements in range of motion observed following manipulation. Finally, they were instructed to do all activities that did not increase their symptoms and to maintain usual activity level within the limits of pain. Advice to maintain usual activity has been found to assist in recovery from LBP.[21] Patient #1

was also instructed to initiate a previously prescribed regime of lumbar spine stabilization exercises to address suspected lumbar spinal instability that may be contributing to his LBP.[230] Patient #2 was also treated with manual distraction mobilization of the right hip and contract-relax stretching of the right hip flexors.

223

Both patients returned for a follow-up appointment three days after the baseline examination and manipulative intervention. The pain rating and ODQ were re-assessed at that time. The changes in NPRS and ODQ scores for both patients are pictured in Figure 3 and Figure 4, respectively.

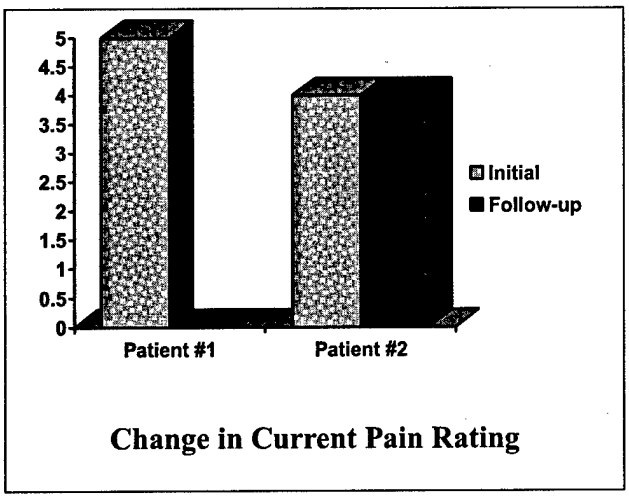


Figure 3. Change in NPRS after three days for both patients. (NPRS scores range from 0 to 10, with 0 being no pain and 10 being maximum pain.)

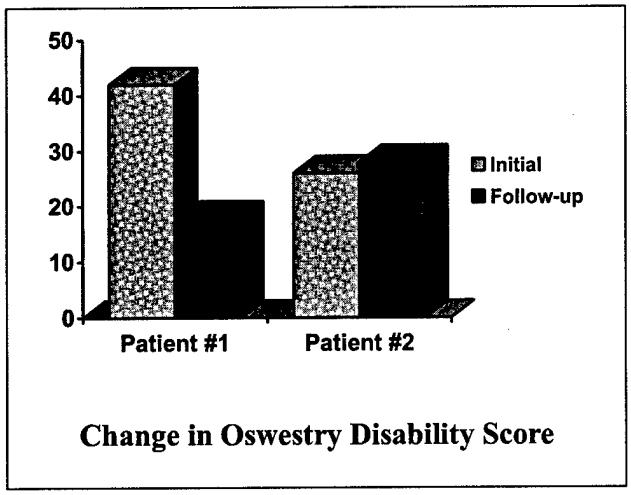


Figure 4. Change in ODQ after three days for both patients (ODQ scores range from 0% to 100% with 0% being no disability and 100% being maximum disability.)

For patient #1, the pain rating decreased from 5/10 to 0/10 and the ODQ decreased from 42% to 18%, a 57% decrease. For patient #2, the pain rating remained unchanged at 4/10, the ODQ was essentially unchanged at 28%, versus 26% at baseline (8% increase).

Patient #1 appeared to have made dramatic improvement following the manipulative intervention and range of motion exercise. After three days he reported no pain, and the 24-point improvement in the ODQ is equivalent to four times the MCID of 6 points that has been established for this instrument.[210] Patient #2 did not appear to benefit from the manipulative intervention and range of motion exercise. His pain rating and ODQ scores were unaffected by the intervention. The first patient did not return for subsequently scheduled visits based on his self-report that he was progressing well with the exercise program and could not coordinate regular physical therapy sessions into his work schedule. He was contacted approximately one month after the initial appointment in which he self-reported that his LBP had continued to improve, and that he was only having symptoms with prolonged standing at work. Ergonomic and appropriate shoe wear recommendations were made, and the patient was discharged from therapy. The second patient was seen for five additional appointments over the ensuing three weeks. Treatment consisted of hip joint mobilizations, flexibility and strengthening exercises for the right hip, and deweighted ambulation on a treadmill. The manipulation was repeated one additional time during the course of care. His status remained unchanged. He was referred for further examination of the right hip. A subsequent MR-arthrogram revealed an anterior labrum and capsular tear of the right hip that eventually required surgical correction.

2.7.4.9.6 Discussion

These two cases highlight the potential value of a CPR for physical therapists. Five clinical findings go into determining a patient's likelihood of success (Table 1). Patient #1 had all five criteria present, indicating a high likelihood of success with the treatment (Table 4). Patient #2 had only one (possibly two) positive findings, indicating a low likelihood of responding to the treatment (Table 4). Both patients were judged to have numerous positive findings that have traditionally suggested the need for treatment of the lumbopelvic region. In these two cases, the CPR more accurately predicted the eventual outcome of the intervention than the traditional lumbopelvic region tests.

The problems with traditional lumbopelvic region tests have been previously discussed at length. However, despite these concerns, many clinicians have continued to rely on traditional tests for dysfunction in the lumbopelvic region.[62,166,189,231,232] This is understandable given that the treatment techniques appear to work for a large number of patients, and a more reasonable alternative to decision-making has not been available. The development of CPRs that are based on an examination of data instead of anatomical and biomechanical theories may offer an alternative that helps clinicians become more efficient and effective practitioners.

Patient #2 illustrates a different advantage of the CPR. Based on traditional clinical reasoning, this patient's pain diagram and physical examination were highly suggestive that spinal manipulation may be an effective intervention. This turned out not to be the case, and in fact, the patient's symptoms were originating from pathology at the hip. Had the CPR been used as the foundation for decision-making for this patient instead of traditional theories, the lack of benefit from spinal manipulation could have been predicted. The value of CPRs is not just their potential

ability to identify patients likely to benefit from a particular intervention, but also their ability to identify patients for whom an alternative course of treatment is more appropriate.

2.7.5 The Second Step: Validating the Clinical Prediction Rule

2.7.5.1 Reasons a Validation Study Might Fail to Support Initial Findings

Although the initial findings based on the development of the CPR may be exciting,[12] McGinn et al[195] have suggested there is a three-step process for developing and testing a CPR prior to promotion for wide-spread implementation of the rule in clinical practice. The first step is to create the CPR. This was the purpose of the initial study to develop the CPR.[12] The second step is to validate the CPR. This step is important to insure the results found in the initial study can be validated in another population.[195,226] There are three potential reasons why a validation study may not support initial findings in the development of a CPR.[233]

First, it is possible that some of the predictors may have occurred by chance.[233] Sackett et al[226] refer to the initial sample in which prognostic factors are identified as the "training set" or "derivation set." This is because the strategy of identifying prognostic factors for an outcome of interest only involves prediction of the outcome, paying no attention to whether the factors identified are even biologically plausible. It is possible the might reveal some "biologically nonsensical and random, non-causal quirk" to be predictive of a given outcome.[226] The second reason is that the predictors identified in the initial development of the CPR may be unique to that sample or the clinicians who participated in the study.[233] Other factors related to the design of the initial study could also influence the outcome.[233] The third reason is that different clinicians in a validation study may fail to accurately apply the CPR or perform the tests and

measures in the CPR differently than in the initial study. For each of these reasons, it is not uncommon to find validation studies that fail to support the initial findings of rigorously developed CPRs.[233]

2.7.5.2 Face Validity of the Criteria in the Spinal Manipulation Clinical Prediction Rule

On the surface, the factors identified in the initial study[12] are at least not obviously spurious and seem to have some face validity for their being predictors of success with manipulation. Intuitively, it seems logical that patients who will benefit from spinal manipulation have low fear-avoidance scores, have symptoms that are relatively acute in nature, exhibit symptoms that do not extend distal to the knee, and exhibit some degree of stiffness in the lumbar spine. A few RCTs have found manipulation to be more beneficial for a subgroup of patients with more acute symptoms,[80,84] and clinical practice guidelines[21,142,143,146] currently recommend manipulation for patients with acute LBP of less than four to 6 weeks duration. The presence of hypomobility in the lumbar spine is also intuitively attractive because it implies that patients with the presence of hypomobility somewhere in the lumbar spine seem to benefit from spinal manipulation. However, this finding must be interpreted cautiously given the lack of reliability for this particular test. The reliability of the therapist's judgment that some hypomobility was or was not present somewhere in the lumbar spine demonstrated a kappa value of 0.13.[234] Reliability at the individual segmental level in the lumbar spine ranged from 0.03 – 0.50.[234] The percent agreement between the examiners was high (78%), and it is likely that the high prevalence of positive findings (85%) in the sample deflated the kappa coefficient.[235] Further work is needed to improve the inter-rater agreement on judgments of segmental mobility before these results can be generalized more broadly.

The only criterion in the CPR that does not appear to be entirely logical is that the presence of at least 35° of hip internal rotation range of motion in at least one hip is associated with success with manipulation. Perhaps increased stiffness in the lumbopelvic spine is associated with a compensatory increase in hip internal rotation range of motion. Previous research has suggested an association between limited hip rotation range of motion and LBP.[236-240] Some authors have speculated that patients with a unilateral restriction in internal rotation range of motion represent a unique pattern that may benefit from a specific manipulative intervention.[240-242] Cibulka et al[240] reported that patients with dysfunction in lumbopelvic region tended to have greater external than internal rotation range of motion on the symptomatic side; however, no relationship was established between this finding and outcome from intervention. Fritz et al[239] recently identified several variables related to hip rotation range of motion that were significantly associated with outcome from an intervention, specifically the failure to improve with spinal manipulation. In general, patients with restricted internal and external range of motion tended not to improve, but more specifically, they demonstrated less side-to-side discrepancy in hip internal rotation range of motion (i.e. more symmetrical) and overall decreased total rotation range of motion. Conversely, patients who tended to improve with spinal manipulation exhibited an increased side-to-side discrepancy in internal rotation range of motion and overall increased internal and external rotation range of motion. Further research is needed to explore the relationship between hip rotation range of motion and outcome from spinal manipulation.

2.8 Purpose

Although the criteria in the CPR seem to have adequate face validity, there is no guarantee that these factors will persist in a different group of patients, even ones with similar characteristics as those used in this initial exploratory study. Consequently, it is entirely possible that some or all

of the initial criteria identified in the initial study[12] were actually spurious in nature, thus negating the clinical utility of the CPR. Therefore, to insure the validity of the initial findings, Sackett et al[226] recommends attempting to validate the initial prognostic factors in a "test set" or "validation set."

Without a control group that did not receive spinal manipulation, one could also question the validity of the CPR developed in the initial study[12] in predicting patients likely to likely benefit from spinal manipulation. Primarily, one could argue that the criteria in the CPR may in fact only identify patients likely to have a favorable natural history of LBP, regardless of the treatment provided. As previously discussed, they attempted to address this concern by establishing a relatively high threshold for determining success, a level that would unlikely be attributable to the favorable natural history of LBP. However, in the absence of a control group, the critic could still argue that any number of treatments (or even no treatment at all) could have been substituted for manipulation, and the same criteria would have surfaced.

2.8.1 Importance of a Validation Study

McGinn et al[243] have established a hierarchy of evidence for CPRs. Without a validation study, the CPR presently corresponds to a lower level, Level IV CPR, which means that further study be performed before a recommendation can be made to apply the CPR clinically. If patients classified as positive on the CPR and receive spinal manipulation achieve an improved outcome compared to patients classified as negative on the CPR but receive spinal manipulation, and compared to patients classified as positive on the CPR but receive a competing treatment such as a stabilization exercise intervention, this would add a substantial margin of validity to the CPR. Specifically, these results will clarify that that the CPR indeed predicts patients likely to likely

benefit from spinal manipulation, rather than merely predict patients who have a favorable natural history of LBP.

To our knowledge, the spinal manipulation CPR is the only one currently reported in the literature to predict outcome from treatment. The results of this study will provide useful information for both clinicians and researchers in physical therapy. First, clinicians will benefit from an easy-to-use CPR that will aid decision-making and may improve outcomes for patients with LBP. If the effect size among patients classified as positive on the CPR and receive spinal manipulation is greater than the effect size among all patients who receive manipulation, this study will be among the first to demonstrate that the power of clinical research for patients with LBP can be improved if patients are classified prior to the intervention. Several researchers have hypothesized that maximizing homogeneity of the sample using classification principles would improve statistical power and provide a better likelihood of identifying evidence for the effectiveness of an intervention, but this study will be the first to test the hypothesis.

Importantly, validation of the CPR will enable a shift up the evidence hierarchy to a Level II CPR,[243] giving clinicians increase confidence in the ability to accurately apply the CPR in a broad spectrum of patients with LBP.

Ultimately, if the CPR is validated in this study, the simplicity of this system should encourage many clinicians to use the CPR and incorporate spinal manipulation as a routine part of their clinical practice. If our hypothesis is supported, these results will also challenge those who persistently want to teach manipulation as an advanced skill wrapped up in a complicated

diagnostic scheme compared to the use of one general technique using a more simple, yet effective CPR. The spinal manipulation CPR appears to be an efficient and practical, evidence-based tool that can be applied by even the novice physical therapist who is familiar with the CPR and the technique that was used in its development. Based on the fact that it only takes approximately five to ten minutes to assess a patient's status with respect to the CPR, clinicians would be able to easily apply the rule to many patients in a busy clinical setting.

2.8.2 Purpose Statement

Therefore, the purpose of this study was to accomplish the second step in the development of a CPR and validate a CPR to identify patients with LBP likely to benefit from spinal manipulation in a multicenter RCT. Multiple therapists and clinical sites in a variety of healthcare settings and geographical regions in the United States were used to better assess the generalizability of the rule.

3. Research Hypotheses

3.1 Specific Aim 1

Determine the validity of a CPR to identify patients with LBP likely to benefit from spinal manipulation.

3.1.1 Hypothesis Aim 1

It was hypothesized that a significant three-way CPR*Intervention*Time interaction would exist to support the notion that outcome from manipulation depends on a patient's status with respect to the CPR. Specifically, it was hypothesized that patients classified as positive on the CPR (i.e.

at least 4/5 criteria met) and received spinal manipulation would experience greater improvement in one- and four-week outcomes compared to patients classified as negative on the CPR and received spinal manipulation, and compared to patients classified as positive on the CPR but received a competing stabilization exercise intervention. Alternatively, it was hypothesized that if the CPR is predicting patients likely to benefit from spinal manipulation, a patient's status should not be able to distinguish between patients who benefit from the stabilization exercise intervention. Therefore, it was further hypothesized that no difference in outcome from the stabilization exercise intervention would exist based on the patient's status with respect to the CPR.

3.2 Specific Aim 2

Determine the effectiveness of spinal manipulation, regardless of the patient's status with respect to the CPR.

3.2.1 Hypothesis Aim 2

It was hypothesized that among all patients in the study, those who received spinal manipulation would achieve greater improvement in one- and four-week outcomes compared to patients who did not receive spinal manipulation, regardless of the patient's status with respect to the CPR. This aim would only be examined if a significant three-way CPR*Intervention*Time interaction from Specific Aim 1 did not exist.

3.3 Specific Aim 3

Compare the treatment effect for spinal manipulation between patients classified as positive on the CPR (i.e. a homogeneous group) versus all patients with LBP (i.e. a heterogeneous group).

3.3.1 Hypothesis Aim 3

It was hypothesized that the effect size in the more homogeneous group of patients classified as positive on the CPR would be larger for the one- and four-week outcomes compared to a heterogeneous group of patients that ignores the CPR. Additionally, it was hypothesized that a lower number needed to treat (NNT) would be observed for patients classified as positive on the CPR compared to all patients in the study and compared to patients classified as negative on the CPR.

4. Research Design and Methods

4.1 Research Design

This project was a RCT to investigate the validity of a CPR to identify patients with LBP likely to likely benefit from spinal manipulation. Patients who met the inclusion criteria and consented to the study completed several self-report measures related to pain, function and disability, and fear-avoidance behaviors. Patients then received a standardized history and physical examination. Upon completion of the clinical examination, study participants were randomly assigned to receive spinal manipulation plus a stabilization exercise intervention or to receive a stabilization exercise intervention alone. Patients were then classified post priori by an examiner blinded to group assignment as to whether they met at least 4/5 criteria in the CPR developed in the initial study.[12] The primary outcome measure was the one-week ODQ score to assess function and disability mirroring the follow-up used in the initial study that developed the CPR.[12] Function and disability was also assessed after four weeks to examine whether any treatment effect was maintained over the duration of the patient's participation in the study. The one- and four-week pain rating was used as a secondary outcome measure.

The independent and dependent variables are outlined in Table 5.

Table 5. Independent and dependent variables in the study.

Independent Variables	**Levels**	**Dependent Variables**
CPR	1. +CPR (at least 4/5) 2. -CPR (less than 4/5)	1. ODQ score 2. NPRS score
Intervention	1. Manipulation Group 2. Exercise Group	
Time (repeated measures factor)	1. Baseline 2. One week 3. Four weeks	

4.2 Methods

4.2.1 Patient Recruitment

Consecutive patients referred to physical therapy for evaluation and treatment of LBP were considered for study participation. A total of 13 physical therapists recruited patients from the following 8 clinical sites:

1. Wilford Hall Medical Center, Lackland AFB (San Antonio, TX)
2. Malcolm Grow Medical Center, Andrews AFB (Washington DC)
3. Wright-Patterson Medical Center, Wright-Patterson AFB (Dayton, OH)
4. Eglin Hospital, Eglin AFB (Fort Walton Beach, Florida)
5. Luke Medical Clinic, Luke AFB (Phoenix, AZ)
6. Hill Medical Clinic, Hill AFB (Ogden, Utah)
7. F.E. Warren Medical Clinic, F.E. Warren AFB (Cheyenne, WY)

8. University of Pittsburgh Medical Center Health System's Centers for Rehab Services (Pittsburgh, PA)

The study was approved by each site's Institutional Review Board (IRB) before patient recruitment and data collection began.

4.2.2 Description of Patients

The study included patients with acute and chronic LBP who met the following inclusion/exclusion listed below. A combination of physical examination and self-report measures were used to assess a patient's eligibility according to each criterion. The method by which each criterion were examined to determine a patient's eligibility is indicated next to the criterion in parentheses. All patients provided informed consent before participation in the study. A copy of the screening examination form that was used is included in Appendix C.

The following inclusion criteria were used to determine a patient's eligibility for the study:

1. Chief complaint of pain and/or numbness in the lumbar spine, buttock, and/or lower extremity (baseline Pain Diagram form and/or self-report)
2. ODQ disability score of at least 30 points (baseline ODQ form)
3. Age at least 18 years and less than 60 years (Demographic Information form and/or self-report)

Because disability was used as the primary outcome of interest, it is important to insure a moderate level of disability was present at the inception of the treatment. Thus patients were required to have at least a baseline ODQ score of 30%. Previous work has shown an average ODQ score of approximately 40% for new patients referred to physical therapy, with a standard

deviation of about 10 points.[225] These values are also similar to those reported in other studies.[216] A minimum ODQ score of 30% allows for the inclusion of a spectrum of patients, but prevents a floor effect from occurring due to low baseline disability scores. Requiring a minimum ODQ score of 30% is also consistent with the inclusion criteria used in the initial study.[12] This insures that patients in this validation study are similar to those used in the initial study that developed the CPR. Also similar to the initial study,[12] patients in this study were not excluded based on the presence of lower extremity symptoms because studies have shown that the lumbopelvic region is capable of referring pain into the lower extremity, even extending distal to the knee.[25,244] These inclusion criteria helped to insure that the sample used in this study was consistent with the patient population for whom the CPR was developed (i.e. patients with LBP with or without lower extremity symptoms with at least a moderate level of disability.) Clinicians do not commonly perform spinal manipulation in individuals under the age of 18, and adults over age 60 with LBP are more likely to have degenerative or stenotic conditions[245] in which manipulation may be contraindicated, thus these two populations were excluded from this study.

The following exclusion criteria were used to determine a patient's ineligibility for the study:

1. Red flags noted in the patient's general medical screening questionnaire (i.e. tumor, spinal compression fracture, metabolic diseases, RA, osteoporosis, prolonged history of steroid use, etc.)
2. Signs consistent with nerve root compression, including <u>any one</u> of the following:
    a. Reproduction of low back or leg pain with straight leg raise at less than 45^0
    b. Muscle weakness involving a major muscle group of the lower extremity
    c. Diminished lower extremity muscle stretch reflex (Quadriceps or Achilles tendon)

d. Diminished or absent sensation to pinprick in any lower extremity dermatome

3. Prior surgery to the lumbar spine or buttock (Demographic Information form and/or self-report)
4. Current pregnancy
5. Inability to comply with treatment schedule (weekly for four weeks)

These criteria were designed to exclude individuals for whom manipulation is contraindicated. In addition, patients had to be able to comply with the four-week treatment schedule. Once patients were admitted to the study, intention-to-treat principles were used, and no patient was removed for non-compliance. However, patients were excluded if they knew ahead of time that they would be unable to comply with the treatment schedule (i.e. traveling extensively during the four-week time period). No individuals were excluded on the basis of gender, race, creed, color, or national or ethnic origin. Therapists recorded the reason for each patient who was ineligible on an eligibility tracking form (Appendix D).

4.2.3 Therapists

Each of the 8 clinical sites had a site coordinator and one or two additional licensed therapists who were trained in the study procedures by one of the investigators. The training session included instruction in the administrative aspects of the study (i.e. informed consent, data collection procedures, etc.) and specific training in the performance of the interventions that were used. The purpose of this training was to insure that the examination and interventions were performed in a similar fashion across sites. The investigator conducting the training individually instructed and observed each therapist in the performance of the manipulative intervention. Each site was provided with a detailed Manual of Standard Operations and Procedures (MSOP) that

outlines all study procedures, to include operational definitions of each physical examination procedure that was used (Appendix E).

4.2.4 Examination Procedures

All eligible patients who consented to participate completed a series of self-report measures, then received a standardized history and physical examination. The self-report measures and physical examination were repeated at the one- and four-week follow-up by an examiner blinded to the patient's status with respect to the CPR.

4.2.4.1 Self-Report Measures

1. Demographic Information (Appendix F) – Demographic information that was collected included age, gender, height, weight, race, employment status, past medical history, and expectation of treatment. Other historical questions that were investigated related to the patient's symptoms included the mechanism of injury, location and nature of the patient's symptoms, number of days since onset, number of previous episodes of LBP, treatment for previous episodes, etc. This information was only collected during the baseline examination.
2. Pain Diagram and Rating (Appendix G) –A body diagram was used to assess the distribution of symptoms.[212,246,247] The location of symptoms was categorized as low back, buttock, thigh, and/or leg (distal to knee) using the method described by Werneke et al,[248] who found high inter-rater reliability (kappa = 0.96). An 11-point scale pain rating scale ranging from 0 (no pain) to 10 (worst imaginable pain) was used to assess current pain intensity and the best and worst level of pain during the last 24 hours.[211,249-252] The average of the three ratings was used to represent the overall level of the patient's pain.

3. Fear-Avoidance Beliefs Questionnaire (FABQ) (Appendix H) – The FABQ quantifies the level of fear of pain and beliefs about avoiding activity in patients with LBP.[214] The instrument consists of 16 items subdivided into two subscales, a 5-item Physical Activity subscale (FABQPA) and a 16-item Work subscale (FABQW). The subscales are reflected in the division of the instrument into separate sections. Questions 1-5 make up the FABQPA subscale, and questions 6-16 make up the FABQW subscale. Decision-making using the CPR requires only the FABQW subscale score. However, all items on the questionnaire should be completed since all items were included when the reliability and validity of the scale was initially established. Each item is scored from 0-6, however not all items within each subscale contribute to the score. Four items (# 2, 3, 4, and 5) are scored for the FABQPA subscale, and 7 items (# 6, 7, 9, 10, 11, 12, and 15) are scored for the FABQW subscale. Each scored item within a particular subscale is summed, thus possible scores range from 0-42 and 0-28 for the FABQW and FABQPA subscales, respectively. Higher scores represent increased fear-avoidance beliefs. Each subscale exists as a separate entity, thus there is no overall FABQ score that consists of the sum of the two subscales. Therapists should insure that all scored items are completed as there is no procedure to adjust for incomplete items. Previous studies have found high level of test-retest reliability for the FABQPA (ICC=0.77) and FABQW (ICC = 0.90) subscales.[253] The FABQW subscale has been associated with current and future disability and work loss in patients with chronic[214,254,255] and acute[256,257] LBP.
4. Oswestry Disability Questionnaire (ODQ) (Appendix I) – The ODQ is as a region-specific disability scale for patients with LBP.[215-219] The questionnaire consists of ten items addressing different aspects of function, each scored from 0-5 with higher values

representing greater disability. The ODQ used in this study was modified to improve compliance. The section on sex life was replaced with one regarding employment/home-making. Previous research has demonstrated the modified version to have high levels of reliability, validity and responsiveness.[210]

4.2.4.2 History and Physical Examination

Patients were asked about the duration of the current episode in days, mechanism of injury, location of symptoms, prior episodes of LBP, and the effect of any treatments received for current or past episodes.[25,258] This information is included in Appendix J.

The components of the physical examination pertinent to the CPR (i.e. segmental mobility of the lumbar spine and hip range of motion) are described in detail below. Descriptions of the remainder of the physical examination procedures are included in the MSOP (Appendix E). A copy of the physical examination form that was used is included in Appendix J.

1. Neurologic Screening Examination: All patients were screened for evidence of nerve root compression. Screening included bilateral straight leg raise tests, manual muscle testing of major muscle groups for myotomes from L_1-S_1, pinprick sensation testing of dermatomes from L_1-S_1, and testing the quadriceps and Achilles reflexes.
2. Physical Impairment Index: Waddell et al[259] described a method of evaluating physical impairment in patients with LBP. The index consists of 7 tests; four range of motion tests (total flexion and extension, average side-bending, and average straight leg raise), and three other tests (bilateral active straight leg raise, active sit-up, and spinal tenderness). Each test is scored as positive (1) or negative (0) based on published cut-off values; resulting in a total score from 0-7. Higher numbers represent increased levels of physical impairment. Waddell

et al[259] found the impairment index could be reliably measured (ICC values 0.86- 0.95, and kappa values 0.48-0.60 for individual tests), and was significantly correlated with disability (r=0.51).

3. Segmental mobility of the lumbar spine: Spring testing of the lumbar spine was tested with the patient prone and the neck in neutral rotation. Testing was performed over the spinous processes of the vertebrae and is both a provocation test and a test of segmental mobility.[186,221] The examiner stood at the head or side of the table and placed the hypothenar eminence of the hand (i.e. the Pisiform bone) over the spinous process of the segment to be tested. With the elbow and wrist extended, the examiner applied a gentle but firm, anteriorly-directed pressure on the spinous process. The stiffness at each segment was judged as normal, hypomobile, or hypermobile. Interpretation of whether a segment is hypomobile was based on the examiner's anticipation of what normal mobility would feel like at that level, and compared to the mobility detected in the segment above and below. In addition, pain provocation at each segment was judged as painful or not painful, and if painful, whether the symptoms were local (i.e. under the examiner's hand) or referred (away from the examiner's hand).
4. Lumbar spine active range of motion - Active range of motion of the lumbopelvic spine was tested with the patient standing according to the procedure described by Waddell et al[259]
5. Hip range of motion – Hip range of motion was tested bilaterally with the patient lying prone with the shoes on, and with the cervical spine at the midline. The examiner placed the leg opposite that to be measured in approximately 30° of hip abduction to enable the tested hip to be freely moved into external rotation. The lower extremity of the side to be tested was kept in line with the body, and the knee on that side was flexed to 90° with the ankle in the neutral

position, and the leg in the vertical position. The inclinometer was placed on the distal aspect of the fibula in line with the bone and was zeroed with the leg in the vertical position. Measurement of hip internal rotation and external rotation was recorded at the point in which the pelvis first began to move. Ellison et al[242] reported excellent inter-rater reliability with these procedures (ICC coefficients ranging between 0.95-0.97).

6. Diagnostic Tests for Lumbopelvic Region Dysfunction - Diagnostic tests that have been purported to identify dysfunction in the lumbopelvic region were also assessed. These procedures included tests designed to assess the symmetry of bony landmarks in the static position (i.e. static symmetry tests), tests to assess the symmetry of bony landmarks with movement (i.e. movement symmetry tests), and tests to reproduce symptoms (i.e. provocation tests).[40] (Appendix A and Appendix E)

4.2.5 Blinding

Three of the five criteria in the CPR can be assessed by patient self-report, thus they are not likely susceptible to rater bias. However, the determination of the presence of segmental hypomobility and hip range of motion testing can potentially be susceptible to rater bias. Moreover, it is possible that a clinician's knowledge of whether a patient meets the criteria in the CPR could potentially bias the treatment of that patient. To minimize these biases, therapists participating in the study were not instructed in the criteria related to the CPR. As a result, they were unaware of the patient's status with respect to the criteria in the CPR. Additionally, patients were not randomized until the baseline examination had been completed, adding additional protection against the possibility for bias to occur. The patient's self-reported change in the ODQ through one week served as the reference criterion, thus the outcome was also not subject to rater bias.

4.2.6 Randomization

A random number generator was used to conduct the randomization, and this procedure was conducted prior to the initiation of the study. The randomization was concealed according to the following procedure. The group assignment was recorded on a label affixed to a 3.5 X 5 inch index card. This card was folded in half such that the label with the patient's group assignment was on the inside of the fold. The folded index card was then placed inside the envelope, and the envelope was sealed. This prevented the potential for the therapist holding the envelope up to the light and visualizing the patient's treatment group assignment through a sealed yet transilluminate envelope. Once the baseline examination was completed, the therapist then opened the randomization envelope indicating the patient's treatment group assignment that corresponded to the patient's identification number. Patients were randomly assigned to one of two intervention arms: 1) spinal manipulation plus a stabilization exercise intervention (Manipulation Group) or 2) stabilization exercise intervention alone (Exercise Group). The patient was notified of the group assignment, and the first treatment session was performed.

4.2.7 Intervention Arms

Patients in both intervention arms attended physical therapy twice a week for the first week and then once a week for the next three weeks, for a total of five visits. The first treatment visit was defined as the first visit in which treatment was administered. In most cases, this occurred on the same visit in which the patient was recruited, or it could have occurred on the patient's second visit, depending on time constraints. If the first treatment session was provided on the patient's second visit, this session always occurred with 24-48 hours of the baseline examination to minimize the likelihood that the patient's status could significantly change due to the passage of

time. All patients were instructed to perform their assigned exercise program once daily on the days they did not attend therapy. Patients were provided an exercise instruction booklet outlining the proper performance and frequency of each exercise. The treatment procedures are described below for patients in each arm of the study. Patients who achieved at least a 50% improvement in their ODQ at the one- and/or four-week follow-up were classified as a success. Otherwise, they were classified as a non-success.

4.2.7.1 Manipulation Group

The intervention received by patients in the Manipulation Group only differed from the Exercise Group during the first two treatment sessions (i.e. during the first week). Beginning on the third session, patients in the Manipulation Group completed the same stabilization exercise intervention as patients in the Exercise Group. During the first two sessions, patients in the Manipulation Group received spinal manipulation and a range of motion exercise only. The manipulative intervention was performed first according to the technique used in the initial study.[12] To perform the technique, the patient was supine. The therapist stood opposite the side to be manipulated and passively moved the patient into side-bending towards the side to be manipulated. The patient was asked to interlock the fingers behind the head. The therapist then rotated the patient, and delivered a quick thrust to the anterior superior iliac spine in a posterior and inferior direction (Figure 1) (video clip).

The side to be manipulated was the more symptomatic side based on the patient's self-report. If the patient was unable to specify a more symptomatic side, the therapist selected either side to be manipulated. Although the manipulative intervention is directed towards one side of the pelvis, Cibulka et al[222] found changes in innominate tilt on both sides of the pelvis after the performance

of this manipulation on one side. Therefore, while manipulating the more symptomatic side provides a consistent approach to determining the side to be manipulated first, it is likely that the manipulation will impact both lumbopelvic regions. Thus the decision as to which side to manipulate is essentially random. In clinical practice, if the manipulative intervention is not successful on one side, we attempt to manipulate the opposite side. Although the risk of a serious complication such as cauda equina syndrome is extremely low,[124,173,174,176] all participating therapists received training in this particular manipulative intervention.

Similar to the procedure used in the initial study,[12] after the manipulation was performed, the therapist recorded whether a cavitation (i.e. "a pop") was either heard or felt by the therapist or patient. If a cavitation occurred, the therapist proceeded to instruct the patient in the range of motion exercise. If no cavitation was produced, the patient was repositioned, and the manipulation was attempted again. If still no cavitation occurred, the therapist attempted to manipulate the opposite side. A maximum of two attempts per side was permitted. If no cavitation was produced after the fourth attempt, the therapist proceeded to instruct the patient in the other treatment components.

Following the manipulative intervention, all patients were instructed in a supine pelvic tilt range of motion exercise as described in Appendix K. This was the same exercise used in the study that developed the CPR.[12] Patients were instructed to perform a set of 10 repetitions in the clinic and 10 repetitions of the exercise 3-4 times daily at home on the days in which the patient did not attend therapy. A copy of the treatment form used by the treating therapist is included in Appendix L. Finally, the patient was instructed to maintain usual activity level within the limits

of pain. Advice to maintain usual activity has been found to assist in recovery from LBP.[21,223] Patients were instructed to do all activities that did not increase symptoms and to avoid activities which aggravated their symptoms. For the remaining three treatment session and for their home exercise program, patients in the Manipulation Group performed the same stabilization exercise intervention as patients in the Exercise Group.

4.2.7.2 Exercise Group

The Exercise Group was treated with a low-stress aerobic and strengthening program. The Agency for Healthcare Policy and Research (AHCPR) Clinical Practice Guidelines for Adults with LBP[21] recommends muscle strengthening exercises for patients with acute LBP, and evidence also supports exercise therapy for individuals with chronic LBP.[18] The strengthening program is designed to target the trunk musculature that has been identified as important stabilizers of the spine in the biomechanical literature.[230,260,261] The theoretical rationale for the stabilization exercise intervention is outlined in Appendix M. The intervention itself can be viewed in Appendix N.

The AHCPR Clinical Practice Guidelines for Adults with LBP[21] also recommends low-stress aerobic exercises for patients with acute LBP. Thus in addition to the strengthening program, an aerobic exercise component was also included. Patients began with a goal of 10 minutes of aerobic exercise on either a stationary bike or treadmill at a self-selected pace. Progression of the aerobic component was performed at the therapist's discretion. Patients in the Exercise Group began the low-stress aerobic exercise component, the warm-up exercise, and then performed the stabilization exercise intervention. Progression of the intervention was accomplished using the

exercise goals listed in Appendix M. A copy of the treatment form used by the treating therapist is included in Appendix O.

4.2.7.3 Post Priori Stratification Based on Clinical Prediction Rule

Following completion of the study, a patient's status with respect to the CPR was determined using the results of the baseline physical examination by an examiner blinded to the patient's group assignment. Patients who met at least 4/5 criteria in the CPR were classified as positive. This decision could have been made a priori by the therapist, but the therapist's knowledge of the patient's status with respect to the criteria in the CPR could serve as a potential source of bias in the treatment of the patient. The patient's status with respect to the CPR is clearly the same whether the judgment is made a priori or post priori. Patients who met at least 4/5 criteria were classified as positive on the CPR. Patients who met three or fewer criteria were classified as negative on the CPR.

4.2.8 Data Analysis

Two statistical packages were used to perform the data analyses for this study. SPSS for Windows, Version 10.1 (SPSS Inc., Chicago, IL) was used to calculate descriptive statistics for the groups and to perform the inferential statistical analyses used in this study. Confidence Interval Analysis, Version 2.0 (Trevor Bryant, University of Southampton, UK) was used to calculate the accuracy statistics for the CPR. Descriptive statistics, including frequency counts for categorical variables and measures of central tendency and dispersion for continuous variables were first calculated to summarize the data. Independent t-tests or Mann-Whitney U tests were performed on continuous data as appropriate, and χ^2 tests of independence were performed on categorical data at baseline to detect differences between the groups on key

demographic variables (i.e. age, gender, race, educational level, symptom acuity, etc.), self-report measures (ODQ score, pain rating, etc.), historical (symptoms acuity, extent of distal symptoms, etc.), and physical examination findings (i.e. lumbar spine range of motion). This was done to determine the adequacy of the randomization procedure in evenly distributing these characteristics between the groups. All data were screened to insure they met the assumptions for the inferential statistical analyses described below.

4.2.8.1 Specific Aim 1

Determine the validity of a CPR to identify patients with LBP likely to benefit from spinal manipulation.

4.2.8.1.1 Hypothesis Aim 1

It was hypothesized that a significant three-way CPR*Intervention*Time interaction would exist to support the notion that outcome from manipulation depends on a patient's status with respect to the CPR. Specifically, it was hypothesized that patients classified as positive on the CPR (i.e. at least 4/5 criteria met) and received spinal manipulation would experience greater improvement in one- and four-week outcomes compared to patients classified as negative on the CPR and received spinal manipulation, and compared to patients classified as positive on the CPR but received a competing stabilization exercise intervention. Alternatively, it was hypothesized that if the CPR is predicting patients likely to benefit from spinal manipulation, a patient's status should not be able to distinguish between patients who benefit from the stabilization exercise intervention. Therefore, it was further hypothesized that no difference in outcome from the stabilization exercise intervention would exist based on the patient's status with respect to the CPR.

4.2.8.1.2 Analysis Aim 1

This aim was examined with a three-way, 2X2X3 repeated measures multivariate analysis of variance (MANOVA). The primary and secondary dependent variables are the ODQ and NPRS scores, respectively. The independent variables were 1) Intervention with two levels (Manipulation vs. Exercise Group), 2) CPR with two levels (+CPR vs. -CPR), and Time with three levels (baseline, one-, and four-week follow-up). The hypothesis of interest was the three-way CPR*Intervention*Time interaction. Planned pairwise comparisons of the simple effects of CPR on Intervention were performed for both the ODQ and NPRS scores at the one- and four-week follow-up using the Bonferroni inequality. The Bonferroni procedure controls the overall family-wise α-level to .05, so that the probability of any single comparison being a Type-I error is not greater than .05.[262]

The first comparison was conducted between patients who received spinal manipulation based on their status with respect to the CPR (i.e. +CPR, Manipulation Group vs. -CPR, Manipulation Group). The second comparison was conducted between patients classified as positive on the CPR based on whether they received spinal manipulation or the stabilization exercise intervention alone (i.e. +CPR, Manipulation Group vs. +CPR, Exercise Group). A third comparison was performed between patients who received the stabilization exercise intervention alone based on their status with respect to the CPR (+CPR, Exercise Group vs. -CPR, Exercise Group). Although it is technically unnecessary to test the overall null hypothesis when planned comparisons are used, the MANOVA as previously described will still be performed to illustrate the three-way CPR*Intervention*Time interaction.

4.2.8.2 Specific Aim 2

Determine the effectiveness of spinal manipulation, regardless of the patient's status with respect to the CPR.

4.2.8.2.1 Hypothesis Aim 2

It was hypothesized that among all patients in the study, those who received spinal manipulation would achieve greater improvement in one- and four-week outcomes compared to patients who did not receive spinal manipulation, regardless of the patient's status with respect to the CPR. This aim would only be examined if a significant three-way CPR*Intervention*Time interaction from Specific Aim 1 did not exist.

4.2.8.2.2 Analysis Aim 2

If necessary, this aim would be examined with a two-way, 2X3 repeated measures MANOVA. The primary and secondary dependent variables would be the ODQ and NPRS scores, respectively. The independent variables would be 1) Intervention with two levels (Manipulation vs. Exercise Group) and 2) Time with three levels (baseline, one, and four weeks after treatment). The hypothesis of interest would be the two-way Intervention * Time interaction. The hypothesis would be supported if the Manipulation Group achieved improved outcomes compared to the Exercise Group at the one- and/or four-week follow-up.

4.2.8.3 Specific Aim 3

Compare the treatment effect for spinal manipulation between patients classified as positive on the CPR (i.e. a homogeneous group) vs. all patients with LBP (i.e. a heterogeneous group).

4.2.8.3.1 Hypothesis Aim 3

It was hypothesized that the effect size in the more homogeneous group of patients classified as positive on the CPR would be larger for the one- and four-week outcomes compared to a heterogeneous group of patients that ignores the CPR. Additionally, it was hypothesized that a lower number needed to treat (NNT) would be observed for patients classified as positive on the CPR compared to all patients in the study and compared to patients classified as negative on the CPR.

4.2.8.3.2 Analysis Aim 3

This aim was examined by calculating the standardized effect size and associated 95% confidence interval based on a patient's status with respect to the CPR at the one- and four-week follow-up. An effect size is a standardized measure of change, and is important for the determination of sample size for future clinical studies. The effect size was calculated as the difference in the mean score on the variable of interest at the relevant follow-up divided by the pooled standard deviation between the groups..[263] The following effect sizes and associated 95% confidence intervals were calculated for both the ODQ and NPRS at the one- and four-week follow (8 total calculations).

1. All patients who received spinal manipulation (n=70) vs. all patients who received the stabilization exercise intervention only (n=61).
2. Only patients classified as positive on the CPR and received spinal manipulation (n=23) vs. all patients who received the stabilization exercise intervention only (n=61).

A significant difference was observed to exist if the 95% confidence intervals did not overlap.

To further illustrate the value of classification and establish whether the benefit is worth the effort to use the CPR in clinical practice, the number need to treat (NNT) statistic and associated 95% confidence interval was also calculated among the following three subgroups of patients who received spinal manipulation according to the procedure described by Sackett et al.[226]

1. All patients
2. Patients classified as positive on the CPR
3. Patients classified as negative on the CPR

This analysis was primarily descriptive in nature, serving to illustrate the importance of classification to improve decision-making and to increase the statistical power of research. However, a significant difference was observed to exist if the 95% confidence intervals did not overlap.

4.2.9 Sample Size and Power

The sample size calculation was conducted a priori using SamplePower™, Release 1.2.[264] based on detecting a significant three-way CPR*Intervention*Time interaction in Specific Aim 1 using the four-week ODQ score at an α-level set to 0.05. The study was powered on the interaction because its detection would contribute most significantly to the validity of the CPR. Based on previous research,[4,5,225] a within-cell standard deviation of 15 points on the ODQ, and a correlation between the covariate and dependent variable of 0.30 ($R^2 = 0.09$) was expected. Given these variables, 21 patients per cell were required to detect a moderate effect size (0.30) for the interaction with 80% power using a two-tailed hypothesis at an α-level of 0.05.[265]

Ninety-four patients were required to achieve 21 patients per cell assuming an even distribution of patients with respect to their status on the CPR and a 10% drop-out rate. However, in the initial study,[12] 30% of patients were positive on the CPR (at least 4/5 criteria met), with the other 70% being negative on the CPR (less than four criteria met). To account for this uneven distribution, and assuming the same distribution in the initial study, approximately 140 patients were required to achieve a total of 42 patients classified as positive on the CPR, (30% of 140 total patients = 42 likely responders [21 for each of the two cells with patients positive on the CPR]). A preliminary analysis of the distribution of patients with respect to the CPR was performed after 50 patients had been enrolled, 36% of which were positive on the CPR.

5. Results

Note: This section primarily presents only results, without providing an interpretation of the data. Interpretation of each table is provided in the Discussion section.

131 consecutive patients referred for evaluation and treatment of LBP were recruited from 13 physical therapists across 8 clinical sites in a variety of healthcare settings and geographical regions in the United States from March 2002 through March 2003. A total of 543 patients were screened for study eligibility (Figure 5).

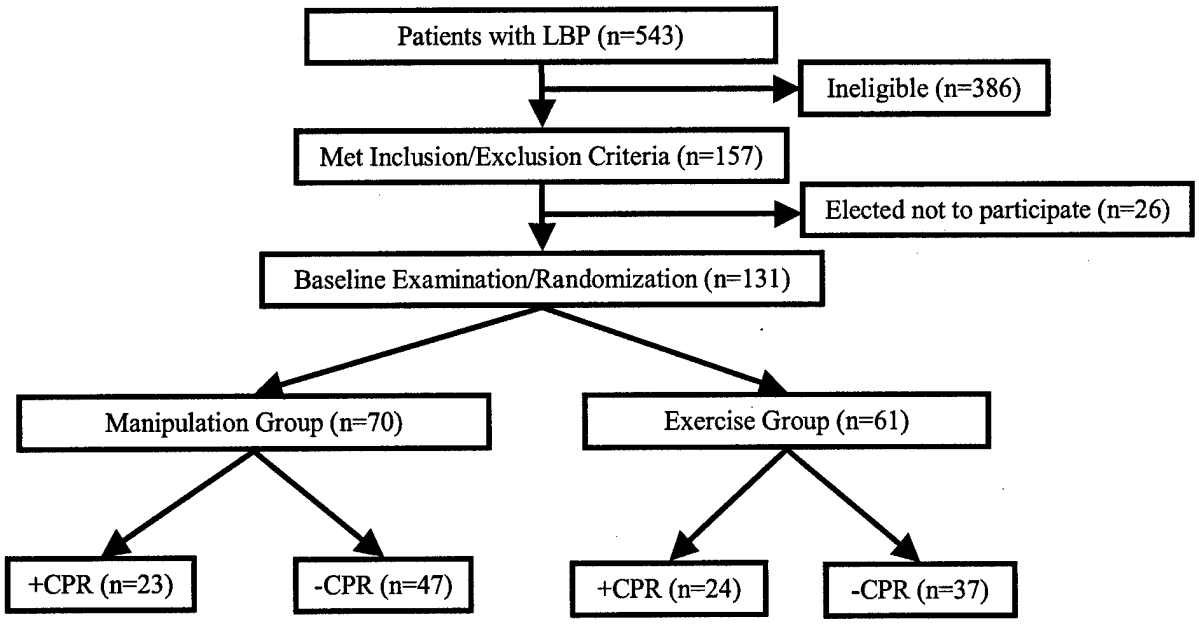


Figure 5. Flow diagram for patient recruitment and randomization.

(Return to p. 98, 98, 98)

Of these patients, 386 patients (71%) were excluded from study participation. The specific reasons for ineligibility and the distribution of patients who were screened for the study are depicted in Figure 6.

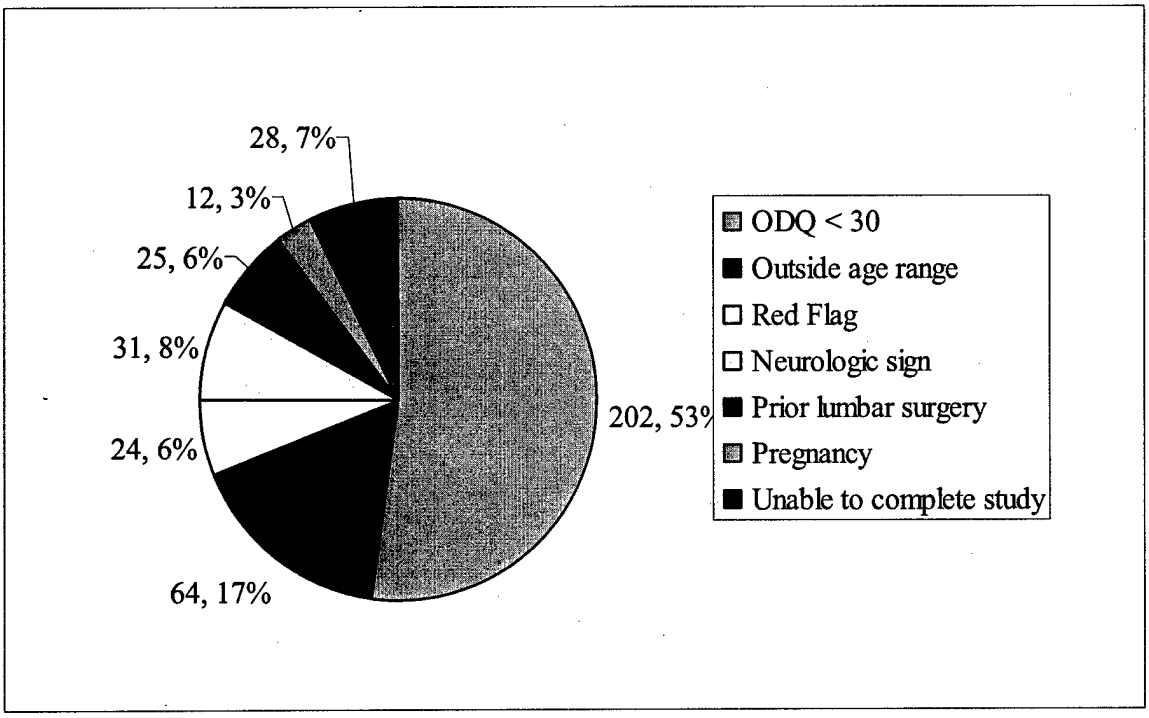


Figure 6. Summary of reasons why patient's were ineligible to participate (n=386).

(Return to p. 98, 149)

The two most common reasons for being excluded were having an ODQ score less than 30% (n=202, 53%) and being outside the specified age range (n=64, 17%) (Figure 6). A total of 157 patients (29%) were deemed eligible to participate, 26 of which elected not to participate (Figure 5). Patients who elected not to participate either did not want to make the commitment of time (n=19) or did not want to be randomized to one of the intervention arms (n=7). The remaining 131 patients provided informed consent and enrolled into the study (Figure 5). Seventy patients were randomized to receive spinal manipulation, and 61 patients were randomized to receive the stabilization exercise intervention (Figure 5). The distribution of patients recruited at each site can be seen in Figure 7.

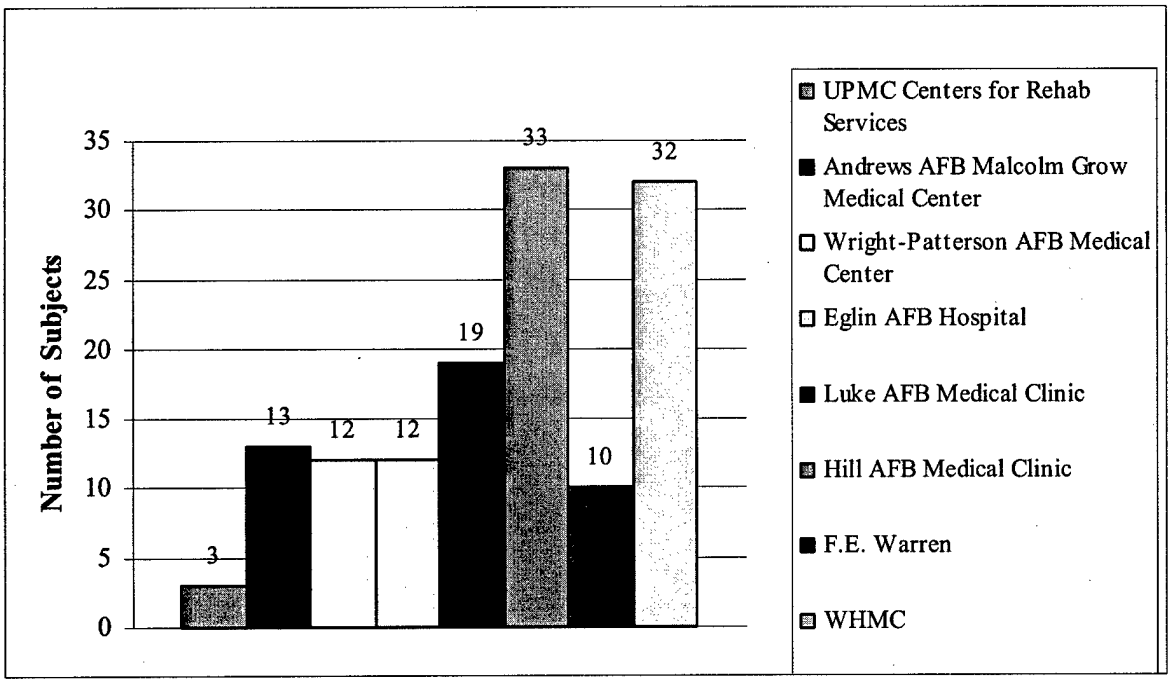


Figure 7. Distribution of patients recruited at each site.

Baseline descriptive statistics for key demographic, self-report measures, historical, and physical examination findings for all patients and within each group are depicted in Table 6.

Table 6. Differences in groups based on key demographic, self-report measures, historical, and physical examination findings. Values represent the mean (SD), except where noted otherwise (when the % sign represents the percentage of patients within the assigned group).

Variable	All Patients (n=131)	Manipulation Group (n=70)	Exercise Group (n=61)	p-value (2-tailed test)
Age (years)	33.9 (11)	33.3 (11)	34.6 (11)	.53
Gender (# of females)	55 (42%)	30 (42.9%)	25 (41.0%)	.83
Race (# of patients)				.27
Caucasian	91 (69.4%)	49 (70.0%)	42 (68.8%)	
African-American	21 (16.0%)	13 (18.6%)	8 (13.1%)	
Hispanic	13 (9.9%)	4 (5.7%)	9 (14.8%)	
Other	6 (4.6%)	4 (5.7%)	2 (3.3%)	
Body Mass Index (kg/m^2)	27.1 (4)	27.7 (5)	26 (4)	.08
Medication use for LBP (# of patients)	110 (84.0%)	61 (87.1%)	49 (80.3%)	.29
Current smoking status (# of patients)	30 (22.9%)	12 (17.1%)	18 (29.5%)	.09
Variable	**All Patients (n=131)**	**Manipulation Group (n=70)**	**Exercise Group (n=61)**	**p-value (2-tailed test)**
Highest level of education completed (# of patients among those who elected to report)				.80
High school	49 (37.4%)	21 (70.0%)	28 (82.3%)	
College (Four-year degree)	10 (7.6%)	6 (20.0%)	4 (11.8%)	
Post-graduate work (At least master's degree)	5 (3.8%)	3 (10.0%)	2 (5.9%)	
Annual household income (# of				.46

patients among those who elected to report)				
Less than $35,000	25 (19.1%)	9 (32.2%)	16 (47.0%)	
$35,000-$70,000	31 (23.7%)	16 (57.1%)	15 (44.2%)	
>$70,000	6 (4.6%	3 (10.7%)	3 (8.8%)	
Variable	**All Patients (n=131)**	**Manipulation Group (n=70)**	**Exercise Group (n=61)**	**p-value (2-tailed test)**
Baseline ODQ score (%)	41.2 (10)	41.4 (10)	40.9 (11)	.77
Baseline NPRS score	5.8 (2)	5.7 (2)	5.9 (2)	.36
Baseline FABQW subscale score	17 (10)	16.5 (10)	17.4 (10)	.63
Variable	**All Patients (n=131)**	**Manipulation Group (n=70)**	**Exercise Group (n=61)**	**p-value (2-tailed test)**
Median number of days for duration of current episode (25th, 75th percentile)	27 (10, 65)	22 (9, 55)	30 (11, 93)	.43
Symptoms distal to the knee (# of patients)	31 (23.7%)	18 (25.7%)	13 (21.3%)	.55
Baseline total flexion range of motion (degrees)	84.2 (28)	84.3 (30)	84.1 (26)	.97
Positive on the CPR (i.e. at least 4/5 criteria met) (# of patients)	47 (35.9%)	23 (32.9%)	24 (39.3)	.440
Number of drop-outs prior to the one-week follow-up	6 (4.6%)	0 (0%)	6 (9.8%)	.007*
Number of drop-outs prior to the four-week follow-up (cumulative)	12 (9.2%)	2 (2.9%)	10 (16.4%)	.007*

(Return to p. 102, 134)

Kolomogorov-Smirnov Z-tests were performed to assess whether continuous data approximated a normal distribution. Except for the duration of the current episode of LBP ($p<.001$), all other continuous variables of interest were found to approximate a normal distribution ($p>.05$). The median number of days and associated 25th and 75th percentile is reported for the duration of current episode of LBP.

An intention-to-treat (ITT) analysis was performed to account for patients who dropped out of the study before the one and four-week follow-up. The last value forward method was used in which the patient's last available score on the ODQ or NPRS was carried forward to the subsequent follow-up. All drop-outs and the specific reason for dropping out of the study are reported in Table 7.

Table 7. Reasons for patients dropping out of study before the one- and four-week follow-up.

	# of patients	
Reason	**One week**	**Four weeks**
Family emergency	1	1
Extended time away from local area	2	2
Excessive time constraints secondary to employment	3	1
Left place of employment (lost to follow-up)	0	1
Developed neurologic signs (+ SLR, myotomal weakness)	0	1
Total	**6**	**6**

(Return to p. 103)

A significantly greater number of patients dropped out of the Exercise Group before both the one- (6 vs. 0 patients) and four-week (four vs. two patients) follow-up ($p=.007$) (Table 6). One

patient in the Exercise Group developed an onset of neurologic signs (positive SLR and myotomal weakness) toward the end of the four-week treatment period and was referred for appropriate management. Upon detailed questioning, it does not appear this onset was related to his participation in the study. The remaining patients clearly dropped out for non study-related reasons, thus this difference in drop-out rate does not appear to be related to the intervention (Table 7). All 131 patients were included in the analysis.

The characteristics of therapists who participated in this study are included in Table 8.

Table 8. Characteristics of participating therapists. Values represent the mean(SD), unless otherwise noted.

n=13	**Mean(SD)**	**Range or Percent**
Age	32.8 (7)	25-47
Gender (# of females)	2 (15.4%)	n/a
Entry-level GPA	3.7 (.28)	3.2-4.0
Years of experience	5.9 (4)	1-16
75-100% of time spent in clinical practice	12 (92.3%)	n/a
n=13	**Mean(SD)**	**Range or Percent**
Highest physical therapy educational degree		n/a
Baccalaureate	3 (23.1%)	
Entry-level master's	8 (61.5%)	
Post-professional master's	2 (15.4%)	
American Board of Physical Therapy Specialties (ABPTS) Orthopaedic Clinical Specialist (OCS) certification	4 (30.8%)	n/a
Residency training	0 (0%)	n/a
n=13	**Mean(SD)**	**Range or Percent**
Years of experience in manual therapy		n/a
< 1 year	3 (23.1%)	
1-5 years	6 (46.2%)	
> 5 years	4 (30.7%)	
Fellow, American Academy of Orthopaedic and Manual Physical Therapists (FAAOMPT)	2 (15.4%)	n/a

Previous experience with the technique used in this study	7 (53.8%)	n/a
Subjective self-rating of manual therapy experience		n/a
Novice/beginner	4 (30.7%)	
Average to above average	9 (69.2%)	
Expert	0 (0%)	

(Return to p. 138, 186)

Table 9 outlines the sources from which participating therapists received their training in spinal manipulation (i.e. high-velocity thrust techniques).

Table 9. Sources from which participating therapists received their training in spinal manipulation (i.e. high-velocity thrust techniques).

Therapists who report receiving at least 20 hours of training in spinal manipulation from the following sources (n=13)	**Number of therapists (%)**
Entry-level education	2 (15.3%)
Continuing education	11 (84.6%)
Pursuit of post-professional physical therapy professional degree	3 (23.1%)
On-the-job training or informal practice time	13 (100%)

(Return to p. 138, 186)

The distribution of patients according to the number of criteria met in the CPR is illustrated in Figure 8.

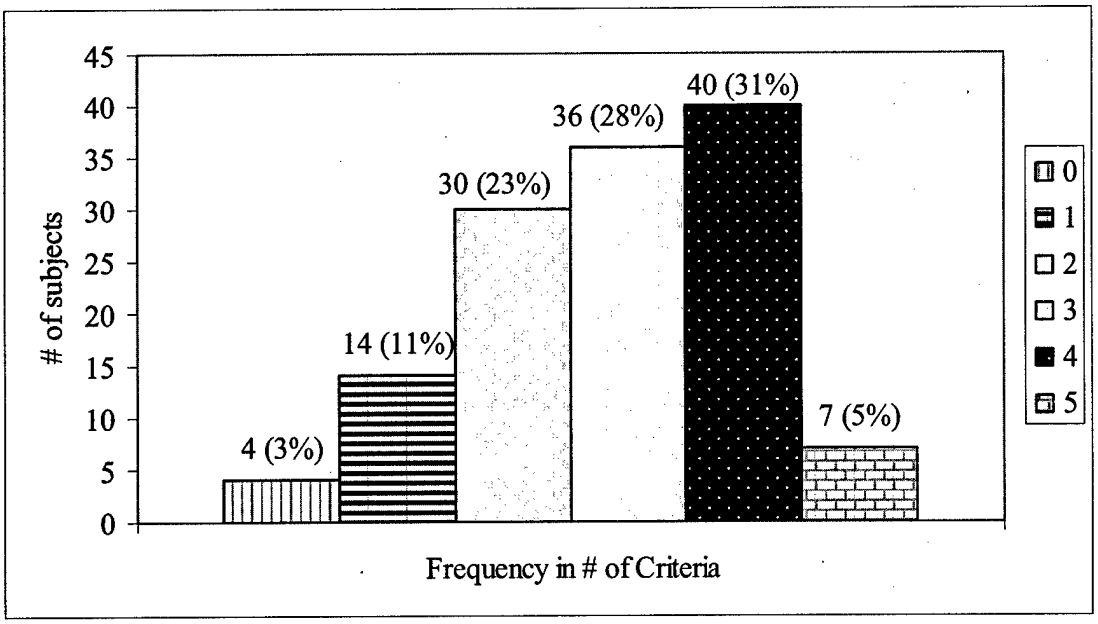


Figure 8. Distribution of patients according to the number of criteria in the CPR met (n=131).

The number of patients in the success and non-success groups at each level of the CPR at the one- and four-week follow-up is depicted in Table 10.

Table 10. Number of patients who received spinal manipulation in the success and non-success groups at each level of the CPR at the one- and four-week follow-up. Success was defined as ⩾ 50% improvement in the ODQ score.

	One week		**Four weeks**	
Number of Predictor Variables Present	**Success**	**Non-Success**	**Success**	**Non-Success**
5	2	0	2	0
4	19	2	20	1
3	8	13	12	9
2	2	16	8	10
1	0	6	2	4
0	0	2	0	2

(Return to p. 139)

Descriptive statistics for the outcome from treatment for patients in both groups is included in Table 11 and Table 12 for the ODQ and NPRS scores, respectively.

Table 11. Descriptive statistics for the raw score, change score, and percent change in ODQ scores at a 2-3 day, one-, and four-week follow-up. Values represent the mean (standard deviation). Positive numbers indicate an improvement in clinical status.

	Baseline	**2-3 days**	**One week**	**Four weeks**	**2-3 day change**	**% change**	**One-week change**	**% change**	**Four-week change**	**% change**
Manipulation Group (n=70)	41.4 (10)	31.2 (14)	23.8 (14)	17.7 (17)	10.2 (13)	24.6%	17.6 (15)	42.5%	23.7 (17)	57.2%
+CPR (n=23)	44.3 (11)	24.6 (15)	13.7 (11)	7.5 (7)	19.7 (14)	44.4%	30.5 (12)	68.8%	36.7 (12)	82.8%
-CPR (n=47)	40.0 (10)	34.4 (12)	28.7 (13)	22.7 (18)	5.6 (9)	14%	11.3 (12)	28.3%	17.3 (15)	43.3%
Exercise Group (n=61)	40.9 (11)	37.2 (12)	33.0 (14)	26.0 (18)	3.7 (12)	9.0%	7.9 (15)	19.3%	14.9 (19)	36.4%
+CPR (n=24)	40.5 (11)	38.1 (13)	34.2 (14)	22.1 (15)	2.4 (13)	5.9%	6.3 (16)	15.6%	18.4 (20)	45.4%
-CPR (n=37)	41.1 (11)	36.7 (12)	32.3 (14)	28.6 (19)	4.5 (11)	10.9%	8.9 (15)	21.7%	12.6 (19)	30.7%

(Return to p. 145)

Table 12. Descriptive statistics for the raw score, change score, and percent change in NPRS scores at the one- and four-week follow-up. Values represent the mean (standard deviation). Positive numbers indicate an improvement in clinical status.

	Baseline	One week	Four weeks	One-week change	% change	Four-week change	% change
Manipulation Group (n=70)	5.7 (2)	3.4 (2)	2.5 (2)	2.3 (2)	40.4%	3.2 (3)	56.1%
+CPR (n=23)	5.9 (2)	2.1 (1)	1.0 (1)	3.8 (2)	64.4%	4.9 (2)	83.1%
-CPR (n=47)	5.5 (2)	4.0 (2)	3.2 (2)	1.6 (2)	29.1%	2.3 (3)	41.8%
Exercise Group (n=61)	5.9 (2)	4.5 (2)	3.6 (2)	1.5 (2)	25.4%	2.3 (2)	39.0%
+CPR (n=24)	6.1 (1)	4.6 (2)	3.4 (2)	1.5 (2)	24.6%	2.7 (2)	44.3%
-CPR (n=37)	5.8 (2)	4.4 (2)	3.8 (2)	1.5 (2)	25.9%	2.0 (2)	34.4%

(Return to p. 145)

5.1 Specific Aim 1

Several methods were used to establish the validity of the CPR. First, data are presented to illustrate the accuracy of the CPR based on the study in which the CPR was originally developed.[12] This can first be illustrated by demonstrating that an increase in the number of criteria a patient met was associated with significant improvement in the ODQ and NPRS at both the one- and four-week follow-up. However, this association did not exist among patients who received the stabilization exercise intervention (Table 13).

Table 13. Association between the number of criteria met at baseline and changes in ODQ and NPRS scores at the one- and four-week follow-up. Values reflect the Pearson correlation coefficient, with positive numbers indicating improved pain and function with an increase in the number of criteria met.

	One week		**Four weeks**	
	Change in ODQ	**Change in NPRS**	**Change in ODQ**	**Change in NPRS**
Manipulation Group (n=70)	.60 (p<.001*)	.48 (p<.001*)	.53 (p<.001*)	.47 (p<.001*)
Exercise Group (n=61)	-.12 (p=.322)	-.066 (p=.615)	.02 (p=.911)	.02 (p=.911)
*significant at p<.05, two-tailed test				

(Return to p. 135, 142)

Table 14 and Table 15 illustrates the accuracy at each level of the CPR to identify patients likely to benefit from spinal manipulation at the one- and four-week follow-up respectively.

Table 14. Accuracy at each level of the CPR among patients who received spinal manipulation at the one-week follow-up. The probability of success is calculated using the positive and negative LR and assumes a pre-test probability of success of 44.3%. Values represent accuracy statistics with 95% confidence intervals for individual variables for predicting success. Success was defined as $\geqslant$ 50% improvement in the ODQ score.

Number of Predictor Variables Present	**Sensitivity**	**Specificity**	**+LR**	**Probability of Success**	**-LR**	**Probability of Success**
All five present	.07 (.02, .21)	1.0 (.91, 1.0)	infinite (.22, infinite)	indeterminate	.93 (.79, 1.08)	42.5%
Four or more present	.68 (.50, .81)	.95 (.83, .99)	13.2 (3.4, 52.1)	91.3%	.34 (.20, .57)	21.3%
Three or more present	.94 (.79, 98)	.62 (.46, .75)	2.4 (1.6, 3.7)	65.6%	.10 (.03, .41)	7.4%
Two or more present	1.0 (.89, 1)	.21 (.11, .36)	1.26 (1, 1.6)	50.0%	0 (0, 1)	approaching 0%
One or more present	1.0 (.89, 1.0)	.05 (.01, .17)	1.05 (.89, 1.2)	45.5%	0 (0, 11)	approaching 0%

(Return to p. 139, 139, 139, 140, 140, 156)

Table 15. Accuracy at each level of the CPR among patients who received spinal manipulation at the four-week follow-up. The probability of success is calculated using the positive and negative LR and assumes a pre-test probability of success of 44.3%. Values represent accuracy statistics with 95% confidence intervals for individual variables for predicting success. Success was defined as $\geqslant$ 50% improvement in the ODQ score.

Number of Predictor Variables Present	**Sensitivity**	**Specificity**	**+LR**	**Probability of Success**	**-LR**	**Probability of Success**
All five present	.05 (.01, .15)	1.0 (.87, 1)	infinite (.38, infinite)	indeterminate	.95 (.85, 1.13)	43.0%
Four or more present	.50 (.36, .64)	.96 (.81, .99)	13.0 (1.9, 90.9)	91.2%	.52 (.38, .71)	29.2%
Three or more present	.77 (.63, .87)	.62 (.43, .78)	2.0 (1.2, 3.4)	61.4%	.37 (.20, .69)	22.7%
Two or more present	.96 (.85, .99)	.23 (.11, .42)	1.2 (1.0, 1.5)	48.%	.20 (.04, .91)	13.7%
One or more present	1.0 (.92, 1.0)	.08 (.02, .24)	1.09 (.94, 1.32)	46.4%	0 (0, 4)	approaching 0%

(Return to p. 139, 140)

Table 16 and Table 17 represent the two-by-two contingency tables used to calculate the accuracy statistics for predicting patients likely to benefit from spinal manipulation at the one- and four-week follow-up, respectively, based on a cut-off of meeting at least 4/5 criteria in the CPR.

Table 16. Accuracy of the CPR to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. A positive CPR was defined as $\geq$ 4/5 criteria met.

	Success	Non-success	Total (%)
+CPR	21	2	23 (32.9%)
-CPR	10	37	47 (67.1%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .68 (.50, .81) Sp: .95 (.83, .99)		+LR: 13.2 (3.4, 52.1) -LR: .34 (.20, .57)	

(Return to p. 134, 136, 138, 138, 144, 144, 154, 154, 155, 159)

Table 17. Accuracy of the CPR to identify patients likely to benefit from spinal manipulation at the four-week follow-up. Success was defined as ⩾ 50% improvement on the ODQ. A positive CPR was defined as ≥ 4/5 criteria met.

	Success	Non-success	Total (%)
+CPR	22	1	23 (32.9%)
-CPR	22	25	47 (67.1%)
Total (%)	44 (62.9%)	26 (37.1%)	70
Sn: .50 (.36, .64) Sp: .96 (.81, .99)		+LR: 13.0 (1.9, 90.9) -LR: .52 (.38, .71)	

(Return to p. 135, 138)

Table 18 and Table 19 represent the two-by-two contingency tables used to calculate the accuracy statistics for predicting patients likely to benefit from spinal manipulation at the one- and four-week follow-up, respectively, based on a cut-off of meeting at least 3/5 criteria in the CPR.

Table 18. Accuracy of the CPR to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. A positive CPR was defined as $\geq$ 3/5 criteria met.

	Success	Non-success	Total (%)
+CPR	29	15	44 (62.9%)
-CPR	2	24	26 (37.1%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .94 (.79, 98) Sp: .62 (.46, .75)		+LR: 2.4 (1.6, 3.7) -LR: .10 (.03, .41)	

(Return to p. 138)

Table 19. Accuracy of the CPR to identify patients likely to benefit from spinal manipulation at the four-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. A positive CPR was defined as $\geq$ 3/5 criteria met.

	Success	Non-success	Total (%)
+CPR	34	10	44 (62.9%)
-CPR	10	16	26 (37.1%)
Total (%)	44 (62.9%)	26 (37.1%)	70
Sn: .77 (.63, .87) Sp: .62 (.43, .78)		+LR: 2.0 (1.2, 3.4) -LR: .37 (.20, .69)	

Table 20 and Table 21 represent the two-by-two contingency table used to calculate the accuracy statistics for predicting patients likely to benefit from the stabilization exercise intervention at the one- and four-week follow-up, respectively, based on a cut-off of meeting at least 4/5 criteria in the CPR.

Table 20. Accuracy of the CPR to identify patients likely to benefit from the stabilization exercise intervention at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. A positive CPR was defined as $\geq$ 4/5 criteria met.

	Success	Non-success	Total (%)
+CPR	3	21	24 (39.3%)
-CPR	4	33	37 (60.7%)
Total (%)	7 (11.5%)	54 (88.5%)	61
Sn: .43 (.16, .75) Sp: .61 (.48, .73)		+LR: 1.1(.44, 2.8) -LR: .94 (.48, 1.8)	

(Return to p. 142)

Table 21. Accuracy of the CPR to identify patients likely to benefit from the stabilization exercise intervention at the four-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. A positive CPR was defined as $\geq$ 4/5 criteria met.

	Success	Non-success	Total (%)
+CPR	10	14	24 (39.3%)
-CPR	12	25	37 (60.7%)
Total (%)	22 (36.0%)	39 (63.9%)	61
Sn: .46 (.27, .65) Sp: .64 (.48, .77)		+LR: 1.3(.68, 2.4) -LR: .85 (.54, 1.3)	

(Return to p. 142)

Table 22 represents a summary of the univariate accuracy for individual items within the CPR to identify patients likely to benefit from spinal manipulation at the one-week follow-up.

Table 22. Summary of the univariate accuracy for individual items within the CPR to identify patients likely to benefit from spinal manipulation at the one-week follow-up.

Number of Predictor Variables Present	**Sensitivity**	**Specificity**	**+LR**	**Probability of Success**	**-LR**	**Probability of Success**
Self-report and history findings						
Duration of current episode of LBP (symptoms<16 days)	.68 (.50, .81)	.85 (.70, .93)	4.4 (2.0, 9.6)	77.8%	.38 (.23, .65)	23.2%
Extent of distal symptoms (no symptoms distal to the knee)	.94 (.79, .98)	.41 (.27, .57)	1.6 (1.2, 2.1)	56.0%	.16 (.04, .63)	11.3%
FABQW subscale score (<19 points)	.52 (.35, .68)	.39 (.25, .54)	.84 (.55, 1.3)	40.1%	1.3 (.73, 2.2)	50.8%
Number of Predictor Variables Present	**Sensitivity**	**Specificity**	**+LR**	**Probability of Success**	**-LR**	**Probability of Success**
Physical Examination findings						
Segmental mobility testing (hypomobility in at least one lumbar spine segment)	.97 (.84, .99)	.49 (.34, .64)	1.9 (1.4, 2.6)	60.2%	.07 (.01, .47)	5.3%
Hip IR range of motion (35° in at least one hip)	.58 (.41, .74)	.69 (.54, .81)	1.9 (1.1, 3.3)	60.2%	.61 (.38, .96)	32.7

(Return to p. 143, 156)

Table 23, Table 24, Table 25, Table 26, and Table 27 represent the two-by-two contingency tables used to calculate the univariate accuracy statistics for individual items within the CPR to identify patients likely to benefit from spinal manipulation.

Table 23. Accuracy of the duration of current episode of LBP to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. Positive was defined as a duration of symptoms < 16 days.

	Success	Non-success	Total (%)
+ duration of symptoms	21	6	27 (38.6%)
- duration of symptoms	10	33	43 (61.4%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .68 (.50, .81) Sp: .85 (.70, .93)		+LR: 4.4 (2.0, 9.6) -LR: .38 (.23, .65)	

(Return to p. 156)

Table 24. Accuracy of the extent of distal symptoms to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. Positive was defined as not having symptoms distal to the knee.

	Success	Non-success	Total (%)
+ extent of distal symptoms	29	23	52 (74.3%)
- extent of distal symptoms	2	16	18 (25.7%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .94 (.79, .98) Sp: .41 (.27, .57)		+LR: 1.6 (1.2, 2.1) -LR: .16 (.04, .63)	

(Return to p. 156)

Table 25. Accuracy of the FABQW subscale score to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as ⩾ 50% improvement on the ODQ. Positive was defined as a FABQW subscale score < 19 points.

	Success	Non-success	Total (%)
+ FABQW subscale score	16	24	40 (57.1%)
- FABQW subscale score	15	15	30 (42.9%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .52 (.35, .68) Sp: .39 (.25, .54)		+LR: .84 (.55, 1.3) -LR: 1.3 (.73, 2.2)	

(Return to p. 143, 156)

Table 26. Accuracy of segmental mobility testing to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as ⩾ 50% improvement on the ODQ. Positive was defined as having at least one hypomobile segment somewhere in the lumbar spine.

	Success	Non-success	Total (%)
+ segmental mobility	30	20	50 (71.4%)
- segmental mobility	1	19	20 (28.6%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .97 (.84, .99) Sp: .49 (.34, .64)		+LR: 1.9 (1.4, 2.6) -LR: .07 (.01, .47)	

(Return to p. 156)

Table 27. Accuracy of hip internal rotation range of motion to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. Positive was defined as a having at least one hip with > 35° of hip internal rotation range of motion.

	Success	Non-success	Total (%)
+ Hip IR range of motion	18	12	30 (42.9%)
- Hip IR range of motion	13	27	40 (57.1%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .58 (.41, .74) Sp: .69 (.54, .81)		+LR: 1.9 (1.1, 3.3) -LR: .61 (.38, .96)	

(Return to p. 156)

In addition to the typical ANOVA assumptions of normality, independence, and homogeneity of variance, Mauchly's test assesses the additional assumption of sphericity associated with a repeated-measures design.[266] This assumption holds that that the variances of differences for all pairs of levels of the repeated-measures factor (i.e. Time in this study) are equal. This assumption was not supported for both the ODQ and NPRS in this study, represented by the significant p-values depicted in Table 28. Unequal variances can result in a 2-3% increase in the probability of committing a Type-I error over that associated with the p-value for the case in which sphericity is assumed.[267]

Table 28. Mauchly's test of sphericity for the repeated measures factor, Time.

Dependent Variable	Mauchly's W	Approximate χ^2	d.f.	p-value	Epsilon - Greenhouse-Geisser	Epsilon - Huynh-Feldt
ODQ	.902	12.976	2	.002	.911	.945
NPRS	.892	14.428	2	.001	.902	.936

The Huynh-Feldt[268] and Greenhouse-Geisser[269] corrections have been proposed to account for a departure from the sphericity assumption. This correction is achieved by increasing the critical F-value by decreasing the degrees of freedom used to determine the critical F-value. Therefore, the calculated F -value must he higher to achieve significance, compensating for the increased probability of committing a Type-I error when the sphericity assumption is violated. It has been shown that the Greenhouse-Geisser epsilon is overly conservative, thus the Huynh-Feldt epsilon appears to provide a more accurate estimate of the actual probability of committing a Type-I error.[268] However, in this case, the results are the same regardless of which correction is used.

The overall three-way CPR*Intervention*Time interaction from the repeated measures MANOVA was significant at an α-level=.05 (p<.001) (Table 29).

Table 29. Summary table of the repeated measures MANOVA for the three-way CPR*Intervention*Time interaction.

Source of Variance	**Value**	**Hypothesis *df***	**Error *df***	**F**	**p-value**
CPR*Intervention*Time	.932	4	506	4.52	<.001*

*significant at α-level=.05 (p-value associated with Wilks' lambda)

(Return to p. 140, 145)

The univariate repeated measures ANOVA for each dependent variable also demonstrates a significant three-way CPR*Intervention*Time interaction for both the ODQ and NPRS in Table 30 and Table 31, respectively. An α-level of .025 was attributed to each of the two dependent variables.

Table 30. Summary table of the univariate repeated measures ANOVA for the three-way CPR*Intervention*Time interaction for the ODQ.

Source of Variance	**SS**	***df***	**MS**	**F**	**p-value**
CPR*Intervention*Time	1808.6	2	904.3	8.54	<.001*
Error	26,908.7	254	105.9		

*significant at α-level=.025 (p-value associated with Huynh-Feldt correction)

(Return to p. 140)

Table 31. Summary table of the univariate repeated ANOVA for the three-way CPR*Intervention*Time interaction for the NPRS.

Source of Variance	**SS**	***df***	**MS**	**F**	**p-value**

CPR*Intervention*Time	21.5	2	10.8	5.16	<.008*
Error	530	254	2.1		

*significant at α-level=.025 (p-value associated with Huynh-Feldt correction)

(Return to p. 140)

After the overall F-test was performed for both the repeated measures MANOVA and the two univariate ANOVAs, the simple effects of interest for the CPR on intervention were analyzed and graphed. Because of the difficulty in visualizing a three-way interaction (which requires a three-dimensional graph), the independent variables of CPR and intervention were collapsed into four groups as follows:

1. +CPR (Manipulation Group)
2. -CPR (Manipulation Group)
3. +CPR (Exercise Group)
4. -CPR (Exercise Group)

This resulted in a plot of four cell means across three points in time, 1) baseline, 2) one-, and 3) four-week follow-up. Figure 9 and Figure 10 represent the plot for the ODQ and NPRS scores, respectively.

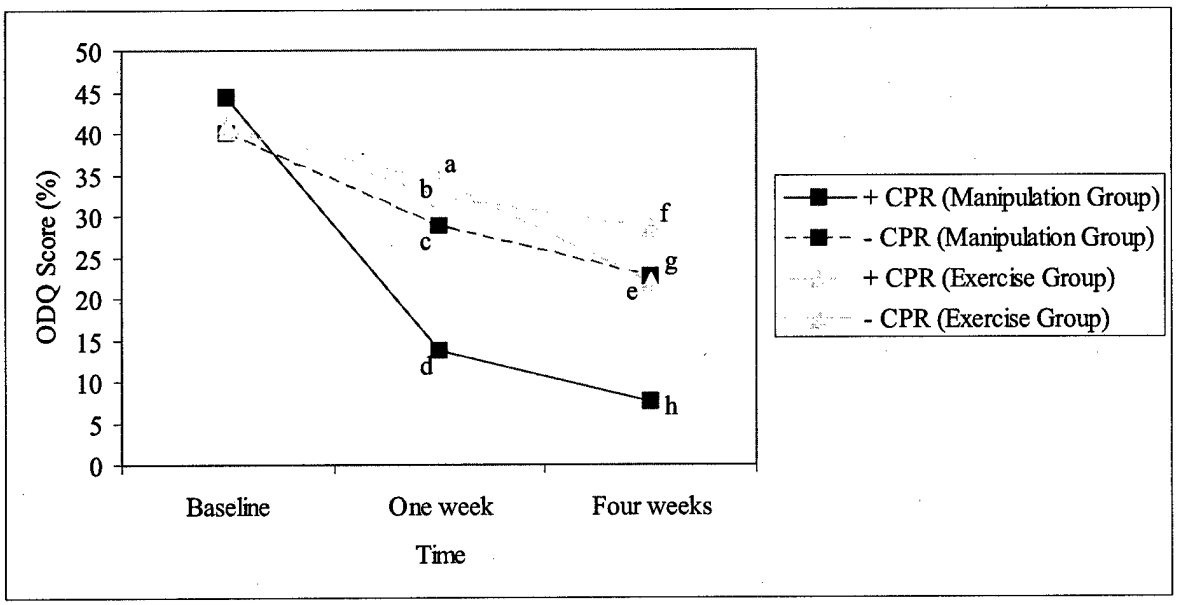


Figure 9. Two-dimensional graphical representation of the three way CPR*Intervention*Time interaction for the ODQ score ($p<.001$).

(Return to p. 127, 141, 142)

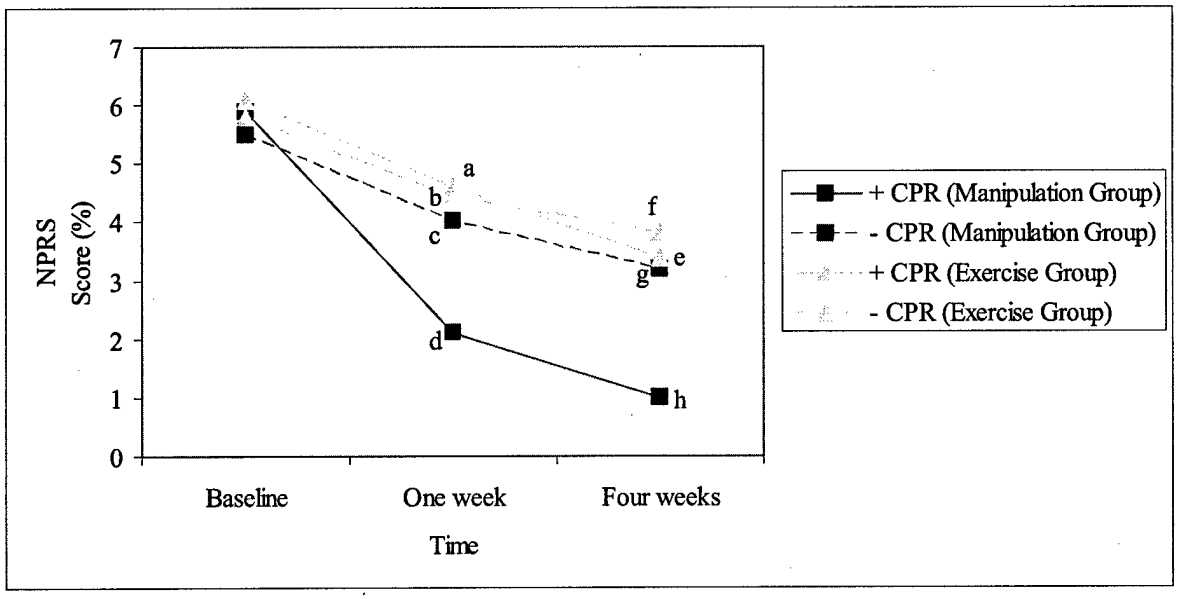


Figure 10. Two-dimensional graphical representation of the three way CPR*Intervention*Time interaction for the NPRS score ($p<.001$).

(Return to p. 127, 141, 142)

The following three comparisons were made at both the one- and four-week follow-up for the ODQ and NPRS scores (Table 32 and Table 33, respectively).

1. +CPR (Manipulation Group) vs. -CPR (Manipulation Group)
2. +CPR (Manipulation Group) vs. +CPR (Exercise Group)
3. +CPR (Exercise Group) vs. -CPR (Exercise Group)

Table 32. Planned pairwise comparisons of the simple effects of CPR on Intervention for the ODQ at the one- and four-week follow-up. The superscripts on the difference scores are depicted in Figure 9.

	One week		**Four weeks**	
	Difference	**p-value**	**Difference**	**p-value**
+CPR (Manipulation Group) vs. -CPR (Manipulation Group)	15.0^{cd} (2.5 MCID**)	<.001*	15.2^{gh} (2.5 MCID**)	<.001*
+CPR (Manipulation Group) vs. +CPR (Exercise Group)	20.4^{ad} (3.4 MCID**)	<.001*	14.6^{eh} (2.4 MCID**)	.003*
+CPR (Exercise Group) vs. -CPR (Exercise Group)	-1.9^{ab} (-.32 MCID**)	.584	6.6^{ef} (3.3 MCID**)	.127

*p-value associated with the Bonferroni inequality significant at family-wise α-level=.05
**The MCID for the ODQ has been demonstrated to be 6%.[210]

(Return to p. 141, 141, 142, 145, 146)

Table 33. Planned pairwise comparisons of the simple effects of CPR on Intervention for the NPRS at the one- and four-week follow-up. The superscripts on the difference scores are depicted in Figure 10.

	One week		**Four weeks**	
	Difference	**p-value**	**Difference**	**p-value**
+CPR (Manipulation Group) vs. -CPR (Manipulation Group)	1.8^{cd} (.9 MCID**)	<.001*	2.1^{gh} (1.1 MCID**)	<.001*
+CPR (Manipulation Group) vs. +CPR (Exercise Group)	2.5^{ad} (1.3 MCID**)	<.001*	2.3^{eh} (1.2 MCID**)	<.001*
+CPR (Exercise Group) vs. -CPR (Exercise Group)	$-.26^{ab}$ (-.13 MCID**)	.622	$.38^{ef}$ (.19 MCID**)	.515

*p-value associated with the Bonferroni inequality, significant at a family-wise α-level=.05
**The MCID for the NPRS has been demonstrated to be two points.[270]

(Return to p. 141, 141, 142, 145, 146)

Although the Bonferroni procedure is a more liberal test in the sense that it gives the researcher credit for planning comparisons in advance,[262] the results are identical to the more conservative Scheffé post hoc multiple comparisons procedure that is commonly used for post hoc (i.e. unplanned) comparisons.

5.2 Specific Aim 2

This aim was not examined because the three-way CPR*Intervention*Time interaction was significant, thus interpretation of the main effect of Intervention is not meaningful. Outcome from manipulation depends upon a patient's status with respect to the CPR.

5.3 Specific Aim 3

Improvement on the ODQ and NPRS for the one-week follow-up based on the patient's status with respect to the spinal manipulation CPR is depicted in Figure 11 and Figure 12, respectively. Improvement at the four-week follow-up is depicted in Figure 13 and Figure 14, respectively.

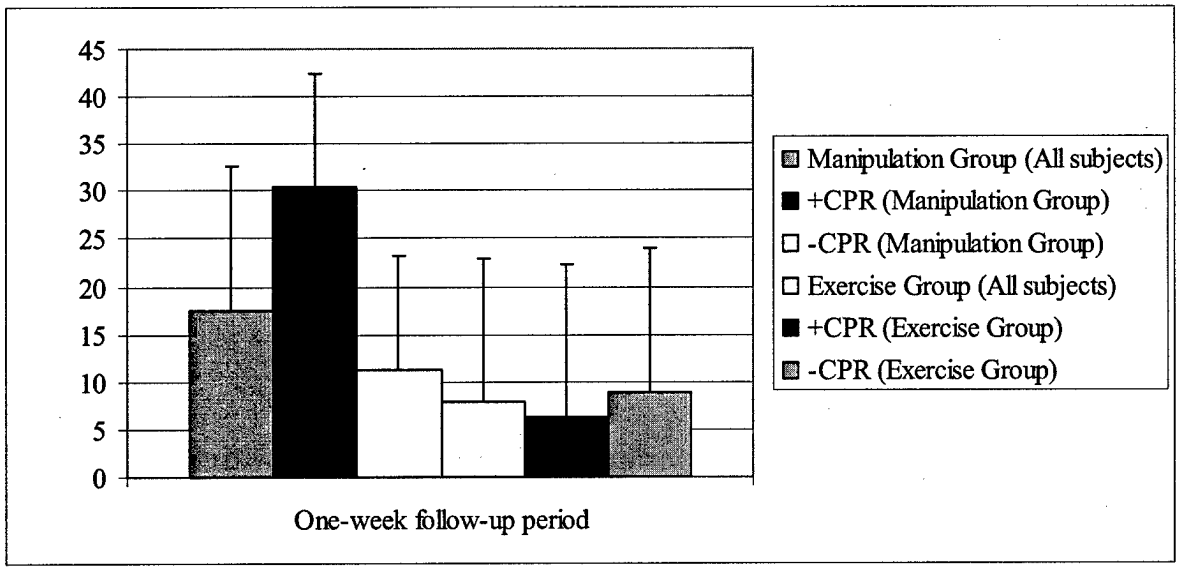


Figure 11. Improvement on the ODQ for the one-week follow-up based on the patient's status with respect to the spinal manipulation CPR. Improvement was defined as the change in disability from baseline to the one-week follow-up ($ODQ_{one\text{-}week}$ - $ODQ_{baseline}$)

(Return to p. 145)

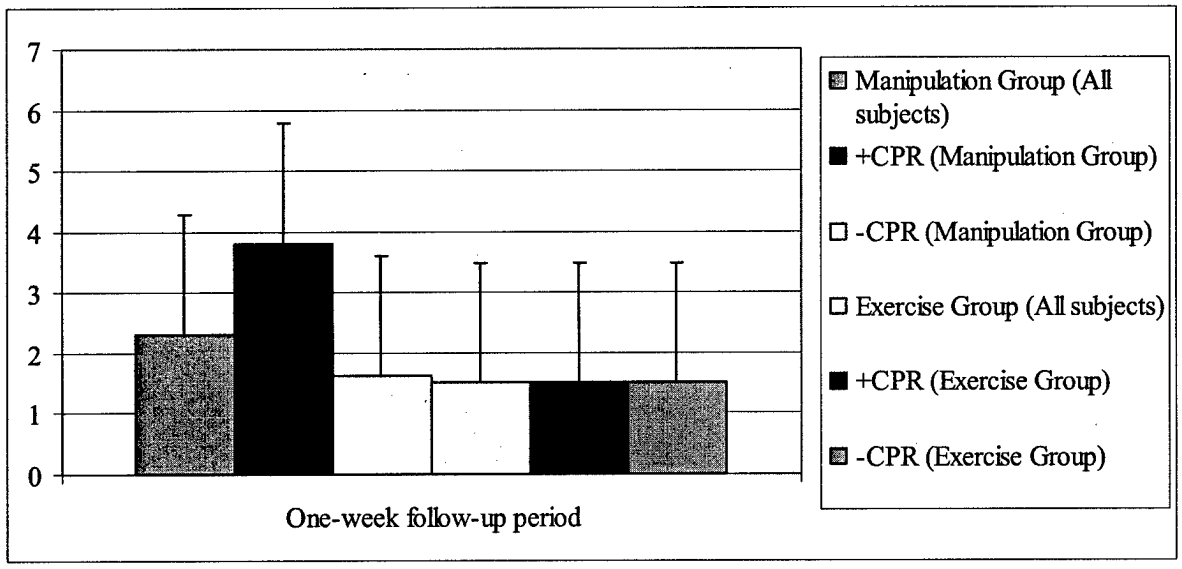


Figure 12. Improvement on the NPRS for the one-week follow-up based on the patient's status with respect to the spinal manipulation CPR. Improvement was defined as the change in pain from baseline to the one-week follow-up ($NPRS_{baseline}$ - $NPRS_{one\text{-}week}$)

(Return to p. 145)

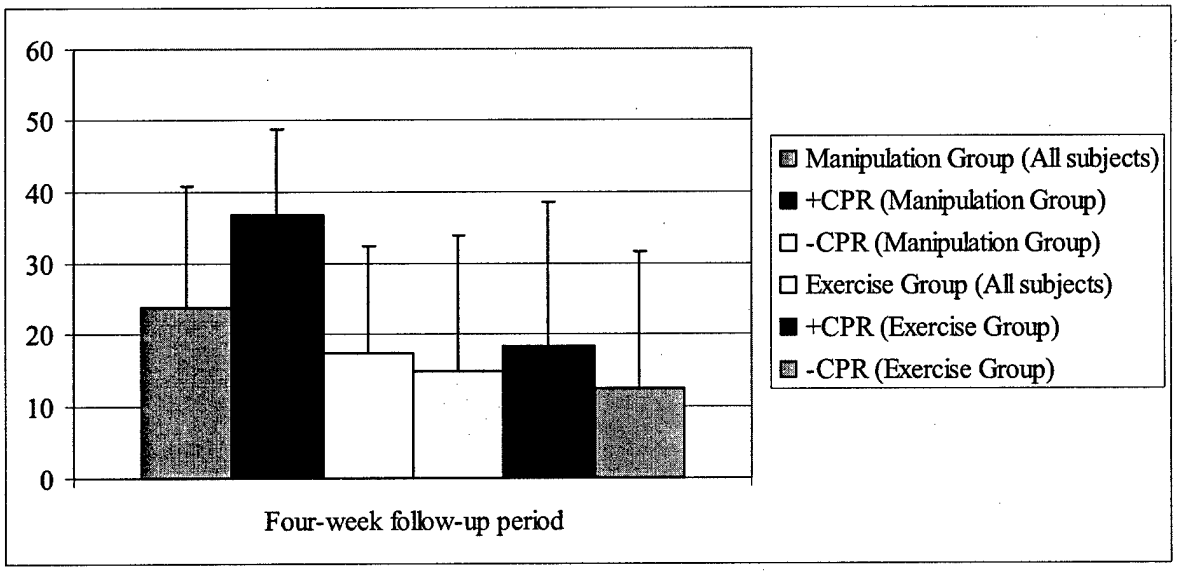


Figure 13. Improvement on the ODQ for the four-week follow-up based on the patient's status with respect to the spinal manipulation CPR. Improvement was defined as the change in disability from baseline to the four-week follow-up ($ODQ_{four\text{-}week}$ - $ODQ_{baseline}$)

(Return to p. 145)

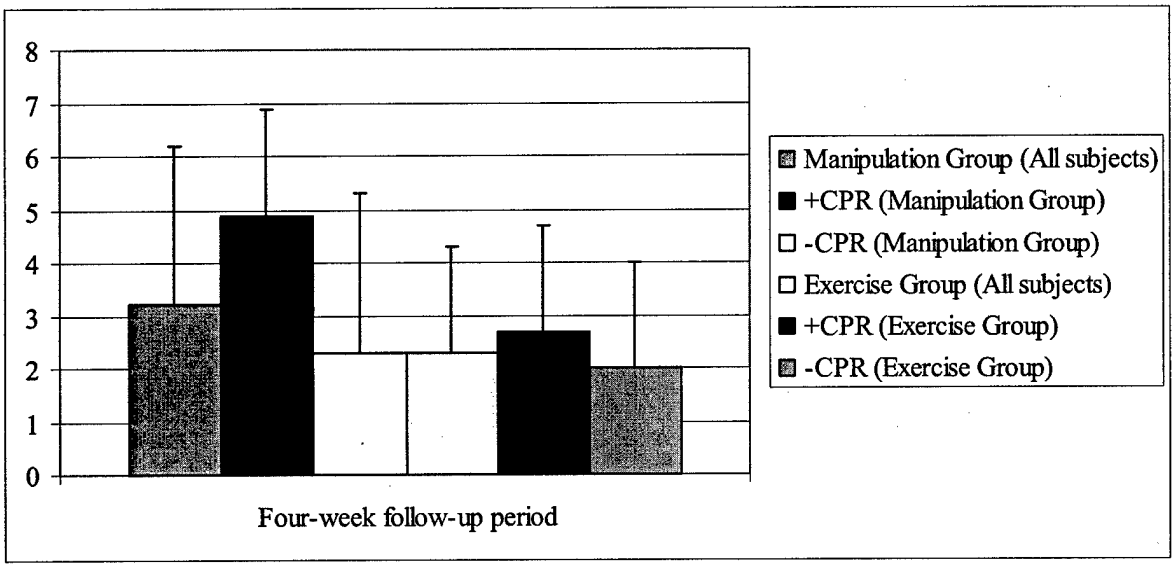


Figure 14. Improvement on the NPRS for the four-week follow-up based on the patient's status with respect to the spinal manipulation CPR. Improvement was defined as the change in pain from baseline to the four-week follow-up ($NPRS_{baseline}$ - $NPRS_{four\text{-}week}$)

(Return to p. 145)

Effect sizes of spinal manipulation for the ODQ and NPRS were calculated at the one- and four-week follow-up and are depicted in Table 34 and Table 35, respectively. An effect size was first calculated as as the difference in final scores at each follow-up period among all patients who received spinal manipulation compared to all patients who received the stabilization exercise intervention. A second effect size was calculated at each follow-up period as the difference in final score among only patients who were positive on the CPR and received spinal manipulation compared to all patients who received the stabilization exercise intervention.

Table 34. Effect size and associated 95% confidence intervals for the ODQ scores at the one- and four-week follow-up. Higher effect sizes represent improvements favoring patients who received spinal manipulation.

	One week	**p-value**	**Four weeks**	**p-value**
All patients (n=70)	.65 (.30, 1.00)	<.001*	.48 (.13, .83)	.01*
+CPR (n=23)	1.45 (.92, 1.97)	<.001*	1.18 (.67, 1.69)	<.001*
*significant at an α-level=.05				

(Return to p. 146, 146, 146)

Table 35. Effect size and associated 95% confidence intervals for the NPRS scores at the one- and four-week follow-up. Higher effect sizes represent improvements favoring patients who received spinal manipulation.

	One week	**p-value**	**Four weeks**	**p-value**
All patients (n=70)	.47 (.13, .82)	.01*	.46 (.12, .81)	.01*
+CPR (n=23)	1.13 (.62, 1.64)	<.001*	1.26 (.74, 1.77)	<.001*
*significant at an α-level=.05				

(Return to p. 146, 146, 146)

The number-needed-to-treat (NNT) statistic based on the patient's status with respect to the CPR is reported in Table 36. This was calculated as the number of patients a clinician must treat with spinal manipulation to avoid one adverse outcome, defined as the patient failing to achieve at least a 50% improvement in the ODQ by the one- and four-week follow-up.

Table 36. NNT based on the patient's status with respect to the CPR. An "adverse" outcome was defined as the patient failing to achieve at least a 50% improvement in the ODQ by the one- and four-week follow-up.

	One-week	**Four-week**
All patients (n=131)	3.1 (2.2, 5.7)	3.7 (2.4, 10.4)
Only +CPR patients (n=47)	1.3 (1.1, 1.9)	1.9 (1.4, 3.5)
Only -CPR patients (n=84)	9.6 (3.9, Infinity)	7.0 (3.0, Infinity)

(Return to p. 183, 183, 183, 183, 183)

6. Discussion

6.1 Random Manipulation of Patients with Low Back Pain

Among all patients in the study, 35.9% (47/131) met at least 4/5 criteria in the CPR (Table 6). This is similar to the results of the initial study that developed the CPR[12] in which 30% (21/71) of patients met at least 4/5 criteria. Among only patients who received spinal manipulation, 44.3% (31/70) achieved at least a 50% improvement in the ODQ at the one-week follow-up, regardless of their status with respect to the CPR (Table 16). In other words, if clinicians were to randomly manipulate patients with non-radicular LBP, they can expect to achieve at least a 50%

improvement in the ODQ within one week approximately 44% of the time. Again, this is similar to the initial study's findings[12] in which 45% (32/71) of patients achieved at least a 50% improvement regardless of their status with respect to the CPR. By the four-week follow-up, 63% of patients (44/70) who received spinal manipulation had experienced at least a 50% improvement in the ODQ (Table 17). Randomly manipulating patients with LBP will result in at least a 50% improvement in the ODQ by the end of one week approximately 45% of the time. Thus it could be reasonably argued that based on the probability of chance alone, an attempt at spinal manipulation is warranted for all patients with non-radicular LBP. A single intervention is rearely used for patients with LBP, thus therapists may recommend other intervention strategies to complement manipulation or for patients who do not improve.

6.2 Accuracy of the Spinal Manipulation Clinical Prediction Rule

Although intuitively attractive because of its simplicity, a random approach does not seem justified in light of evidence for a simple CPR that can improve decision-making and accurately establish a patient's prognosis after receiving spinal manipulation. As a first approximation of the CPR's validity, one can assess the association between the number of criteria met in the CPR and outcome. If the CPR is related to outcome, a linear association between the number of criteria met and outcome from treatment might be suspected. For patients who received spinal manipulation, an increase in the number of criteria was significantly associated with improved pain ($r=.48$, $p<.001$) and function ($r=.60$, $p<.001$) at the one-week follow-up, and a similar finding is observed for the four-week follow-up (Table 13).

However, the primary objective of the CPR is to increase the post-test probability of success sufficiently to influence decision-making. Because this study sought primarily to identify

patients likely to benefit from spinal manipulation, the primary statistic of interest was the positive LR. The positive LR expresses the change in odds favoring the outcome when the patient meets the criteria in the CPR.[226] An accurate CPR should therefore have a large positive LR. According to Jaeschke et al[227] accuracy can be considered moderate when the positive LR is greater than 5.0. Accuracy is substantial when the positive LR is greater than 10.0.[227]

Similar to the findings in the initial study that developed the CPR,[12] a threshold of at least 4/5 criteria met in the CPR maximizes the positive LR in distinguishing between patients who are classified as a success ($\geqslant$50% improvement on the ODQ) and non-success (<50% improvement on the ODQ) with spinal manipulation, yielding a positive LR and 95% confidence interval of 13.2 (3.4, 52.1) (Table 16). Based on a pre-test probability of success of 44.3%, patients who meet at least 4/5 criteria in the CPR and receive spinal manipulation have a 91.3% chance of achieving at least a 50% improvement one-week after treatment, representing an increase in probability on the order of 50% (Figure 15).

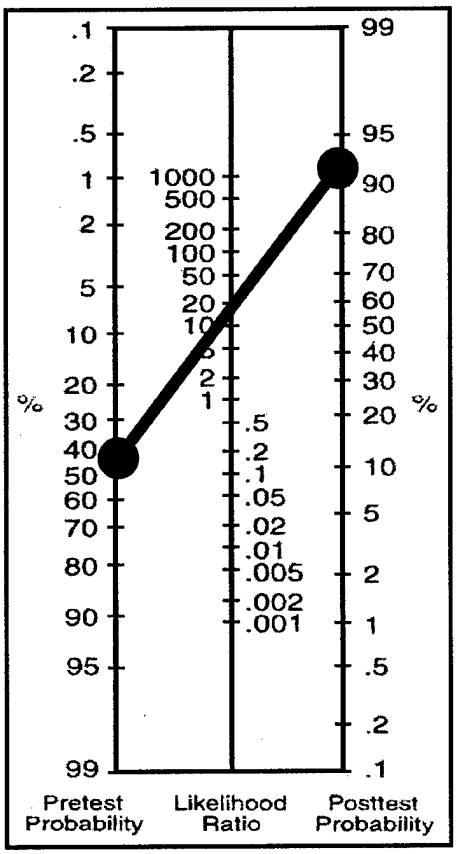


Figure 15. Fagin's nomogram illustrating the shift in post-test probability from 44.3% to 91.2% at the one-week follow-up for patients positive on the CPR who receive spinal manipulation (positive LR=13.2).

A positive LR for 4/5 criteria of 13.2 (Table 16) is smaller than the positive LR of 24.4 (4.63, 139.4) observed in the initial study that developed the CPR.[12] It is possible that the drop in the positive LR in this validation study may partially be attributed to the use of 13 examiners of varying levels of experience (Table 8 and Table 9) distributed across 8 clinical sites, thus likely increasing the overall measurement error. Different therapists may also apply the CPR and perform the manipulative intervention slightly differently. However, the 91.2% post-test probability observed in this study is only slightly smaller than the 95% post-test probability demonstrated in the initial study.[12] More importantly, this level of certainty clearly seems adequate to influence decision-making, and the almost negligible decrease in accuracy seems to be well worth the increased generalizability of the CPR by using multiple examiners and clinical sites. Even if the lower bound of the 95% confidence interval of 3.4 is presumed to be the point estimate (Table 16), the post-test probability of success is 73.0%, which still seems adequate to justify an attempt at spinal manipulation. Similar levels of accuracy were observed at the four-week follow-up (Table 17), supporting the prognostic value of the CPR at a longer follow-up than was initially studied.[12]

When the threshold for defining a positive CPR is reduced to having at least 3/5 criteria met, the positive LR was 2.4 (1.6, 3.7), resulting in a 65.6% post-probability of success at the one-week follow-up (Table 18). This is also similar to the positive LR of 2.6 (1.8, 4.2) (68% probability of success) observed in the study that developed the CPR.[12] Given the ease with which the CPR is applied and manipulative intervention can be performed, and in light in the extremely low risks,[124,173,174,176] an argument can easily be made that this level of probability is still sufficient to justify an attempt at spinal manipulation.

Patients who met fewer than 3/5 criteria are less likely to benefit from spinal manipulation. Similar to the findings in the study that developed the CPR,[12] when less than three criteria were met among patients in this study, the probability of success is essentially no better than the probability of success if you were to randomly manipulate patients with non-radicular LBP (Table 14). Only two patients who met 2/5 criteria achieved at least a 50% improvement in the ODQ at the one-week follow-up, and no patients who met only one (n=6) or none (n=2) of the criteria were classified as a success (Table 10). Thus the clinician may want to consider other interventions with a higher probability of success for these patients. Similar levels of accuracy are observed at the four-week follow-up (Table 15), supporting the prognostic value of the CPR at a longer follow-up than was initially studied.[12]

Although the positive LR was the primary statistic of interest, there is also some value in the interpretation of the negative LR. The negative LR expresses the change in odds favoring the outcome when the patient does not meet the criteria in the CPR.[226] An accurate CPR should therefore have a small negative LR. According to Jaeschke et al[227] accuracy can be considered moderate when the negative LR is less than .20. Accuracy is substantial when the negative LR is less than .10.[227] The negative LR for achieving at least a 50% improvement in the ODQ at the one-week follow-up was .34 (.20, .57) for patients who did not meet at least 4/5 criteria in the CPR (Table 14). Based on a pre-test probability of success of 44.3%, the post-test probability of success for these patients is reduced to 21.3% (Table 14). Although perhaps not a definitive shift in probability, clinicians can be less certain that spinal manipulation will be effective when the patient has less than 4/5 criteria. However, the negative LR becomes more meaningful when

even fewer criteria in the CPR are met. The negative LR was .10 (.03, .41) for patients who met fewer than three criteria in the CPR, reducing the post-test probability of success to only 7.4% (Table 14). This is likely a definitive shift in probability where the clinician is more certain the patient will not likely benefit from spinal manipulation. With fewer than two criteria met, the post-test probability approaches 0%, suggesting that clinicians should consider other interventions for patients who meet none or only one criterion in the CPR (Table 14). Similar negative LRs are observed at the four-week follow-up (Table 15), supporting the prognostic value of the CPR at a longer follow-up than was initially studied.[12]

6.3 Outcome from Spinal Manipulation Depends upon the Clinical Prediction Rule

The significant three-way CPR*Intervention*Time interaction for the overall repeated measures MANOVA ($p<.001$) supports the notion that outcome from spinal manipulation depends upon the patient's status with respect to the CPR (Table 29). The univariate repeated measures ANOVA for the three-way CPR*Intervention*Time interaction was also significant for the ODQ ($p<.001$) (Table 30) and NPRS ($p<.008$) (Table 31). If the spinal manipulation CPR is to be useful for decision-making to identify patients likely to benefit from this intervention, patients classified as positive on the CPR and receive spinal manipulation should demonstrate improved outcomes compared to patients classified as negative on the CPR and receive spinal manipulation, and compared to patients classified as positive on the CPR but receive an alternative approach such as a stabilization exercise intervention. This was in fact the case. Patients classified as positive on the CPR and received spinal manipulation achieved 2.5 times the MCID of 6 points on the ODQ compared to patients classified as negative on the CPR and received spinal manipulation, and 3.4 times the MCID compared to patients classified as positive

on the CPR but received the stabilization exercise intervention alone ($p<.001$) (Figure 9, Table 32). Similar results were observed for the NPRS scores ($p<.001$) (Figure 10, Table 33).

The final outcome assessment in the initial study that developed the CPR[12] was approximately one week. However, the results from this validation study were also maintained at the four-week follow-up for both the ODQ ($p<.003$) and NPRS ($p<.001$) (Table 32 and Table 33, respectively). This suggests the CPR continues to be useful in establishing a patient's prognosis from spinal manipulation beyond a relatively short-term one-week follow-up. Further study will establish the value of the CPR to predict outcome at a 6-month follow-up.

6.3.1 Clinical Prediction Rule Does Not Predict Favorable Natural History

For classification to be meaningful, a CPR that identifies patients likely to benefit from a specific intervention needs to distinguish between patients who may not alternatively benefit from the passage of time, or more importantly, between a competing intervention that also has some evidence for its effectiveness. Because no control group was included in the initial study that developed the CPR,[12] a case could be made that the criteria were merely identifying patients with a favorable natural history. In other words, patients who met the criteria in the CPR may have been likely to benefit from a variety of interventions or the passage of time, thus improvement could not solely be attributed to receiving spinal manipulation. This validation study directly assesses the tenability of this notion because patients in the control group did not receive a sham placebo intervention or no treatment at all. Rather, they completed a legitimate, competing alternative stabilization exercise intervention, which clearly has shown to be effective for patients with LBP.[271]

If the CPR were merely predicting the favorable natural history if LBP, an association between outcome from the stabilization exercise intervention and the number of criteria met in the CPR should exist. However, an association did not exist at either the one- or four-week follow-up (Table 13). More specifically, if the CPR were simply predicting the favorable natural history of LBP, it should accurately distinguish between patients likely to benefit from a variety of interventions, or perhaps from simply the passage of time. In this study, a patient's status with respect to the CPR should accurately identify patients likely to benefit from the stabilization exercise intervention if improvements could be attributed to a favorable natural history. However, this was not the case. The positive LR for identifying patients likely to benefit from the stabilization exercise intervention at the one-week follow-up was only 1.1 (.44, 2.8) (Table 20). A similarly small positive LR of 1.3 (.68, 2.4) exists for prediction of outcome from the stabilization exercise intervention at the four-week follow-up (Table 21). A positive LR close to one suggests small shifts in the post-test probability of success from the intervention that is not useful for decision-making.[227] This was confirmed in the analysis of the simple effects from the three-way ANOVA. There was no difference in outcome on the ODQ at the one- ($p=.584$) or four-week ($p=.127$) follow-up among patients who received the stabilization exercise intervention alone based on the patient's status with respect to the CPR (Figure 9, Table 32). Similar non-significant effects were observed for the NPRS at the one- and four-week follow-up (Figure 10, Table 33). Therefore, the results of this study clearly support the notion that the CPR is identifying patients likely to benefit from spinal manipulation rather than predicting the favorable natural history of LBP.

6.4 The Role of the Fear-Avoidance Beliefs with Spinal Manipulation

Of all the individual criteria in the CPR, the patient's score on the FABQW subscale has the least univariate diagnostic accuracy (Table 22 and Table 25). It seems reasonable to suspect that the overall accuracy of the CPR would increase if this criterion was removed from the CPR. If a modified CPR is developed using the presence of at least 3/4 criteria to qualify as being positive on the CPR, the following two-by-two contingency table is generated for the one-week follow-up (Table 37).

Table 37. Accuracy of a modified CPR to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. A positive CPR was defined as $\geq$ 3/4 criteria met. The FABQW subscale score is excluded.

	Success	Non-success	Total (%)
+CPR	26	6	32 (45.7%)
-CPR	5	33	38 (54.3%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .84 (.67, .93) Sp: .85 (.70, .93)		+LR: 5.5 (2.6, 11.6) -LR: .19 (.08, .43)	

(Return to p. 143, 143, 143, 144)

However, rather than increasing, the positive LR to identify patients likely to benefit from spinal manipulation at the one-week follow-up drops to 5.5 (Table 37), from the original positive LR of 13.2 found when all five criteria are considered (Table 37). Interestingly, ignoring the FABQW subscale score actually results in five additional true positive findings (26 vs. 21), increasing the sensitivity of the CPR to .84 (Table 37), compared to a sensitivity of .68 when the FABQW

subscale score is considered (Table 16). However, this 17% increase in sensitivity comes at the expense of four additional false positive findings (6 vs. two), resulting in a 10% decrease in specificity (.95 to .85) (Table 16 and Table 37, respectively). This 10% decrease in specificity causes the denominator in the calculation of the positive LR (Sn/[1-Sp]) to become larger, resulting in an overall smaller positive LR. In other words, a drop in specificity has a more detrimental impact on the positive LR than does a similar increase in sensitivity. Practically speaking, the four additional patients classified as being a false positive when the FABQW subscale score is ignored all tended to have high FABQW subscale scores. This suggests that patients with high fear-avoidance beliefs generally may not benefit from spinal manipulation. Recent evidence suggests that these patients seem to benefit from a psychosocial treatment approach rather than a more traditional biomedical model.[272] The value of the FABQW subscale appears to be in its ability to identify patients with very high scores as being unlikely to benefit from manipulation, thus contributing to an overall increased accuracy of the CPR, despite its relatively poor diagnostic accuracy when considered in a univariate fashion. Future work from this study can explore whether an upward increase in the cut-off score would be useful to improve the overall accuracy of the CPR, examine the influence of fear-avoidance beliefs on outcome from spinal manipulation, and determine if fear-avoidance beliefs change in response to spinal manipulation.

6.5 Increasing the Power of Clinical Research

It has been suggested that classification will enhance the power of clinical research by permitting researchers to study more homogenous groups of patients;[162,163] however, to our knowledge, this notion has not been explicitly examined. This is important because a more powerful study will

enhance the likelihood of identifying evidence for the effectiveness of an intervention, when a treatment effect might otherwise be masked in a more heterogeneous sample.

6.5.1 Descriptive Illustration of the Value of Classification

Descriptive statistics for the improvements in pain and function can be viewed in Table 11 and Table 12 to illustrate the value of classification. These data are also represented visually in Figure 11, Figure 12, Figure 13, and Figure 14. Improvements in both pain and function were larger among patients classified as positive on the CPR and received spinal manipulation compared to patients classified as negative on the CPR and received spinal manipulation, and compared to patients classified as positive on the CPR and received the stabilization exercise intervention (Table 32 and Table 33). These results illustrate the value of classification to improve decision-making.

6.5.2 Inferential Illustration of the Value of Classification

The existence of a significant three-way CPR*Intervention*Time interaction is perhaps the best inferential illustration of the meaningfulness of classification (Table 29). Based on the significant interaction, it is incorrect to suggest that manipulation in general is better than a spinal stabilization exercise alone. Rather, outcome from spinal manipulation depends upon a patient's status with respect to the CPR, supporting the value of the classification process to improve decision-making.

6.5.3 Illustration of the Value of Classification Using Effect Sizes

An effect size is a standardized measure of change, and is important for the determination of sample size for clinical studies. However, an additional means to illustrate the value of

classification can be observed by comparing the effect size and associated 95% confidence interval among all patients who received spinal manipulation versus only those classified as positive on the CPR. Table 34 depicts the effect size and associated 95% confidence intervals at the one- and four-week follow-up for the ODQ. A similar table is provided for the NPRS (Table 35). When only patients classified as positive on the CPR are considered, the effect size favoring spinal manipulation was twice as large than when the CPR is ignored. Although the effect size itself remains significant for both the ODQ and NPRS even when all patients are included (p=.01) (Table 34 and Table 35), this may not occur when classification for interventions with a smaller effect is examined.

Interestingly, the confidence intervals between the two effect sizes for both the ODQ (Table 34) and the NPRS (Table 35) slightly overlap, thus a definitive statement cannot be made that they're statistically different. However, there is an imbalance in the sample size between the groups (23 vs. 61 patients), which results in a wider confidence interval around the effect size for patients classified as positive on the CPR (n=23) than if the groups were more even. In fact, increasing the number of patients classified as positive on the CPR to 34 is sufficient to avoid any overlap in the confidence intervals for both the ODQ and NPRS. More importantly, the differences in the raw scores are clearly clinically meaningful (Table 32 and Table 33).

The results of this study support the notion that the power of clinical research for patients with LBP can be improved if patients are classified prior to the intervention. The failure to adequately consider the importance of classification is illustrated in the results of a recent systematic review[138] conducted by investigators at the RAND corporation. They concluded that spinal

manipulation does not appear to be more effective than other interventions. Based on the apparently conclusive negative results of this review in their opinion, they further question the need for future clinical trials related to spinal manipulation. This seems to be a dramatic shift in guidance compared to previous results of the authors' own reviews that suggest at least a short-term benefit for spinal manipulation compared to other active interventions.[69,119,122,124,129] Perhaps more importantly, the current review gives scant attention to the notion that a subgroup of patients likely to benefit may exist. The authors conclude the following:

"While not all of the 95% [confidence intervals] in our analysis exclude improvements of moderate clinical importance, most do. We interpret this to mean that spinal manipulative therapy is very unlikely to be a particularly effective therapy for any group of patients with back pain. While it is conceivable that spinal manipulative therapy is very effective for a subgroup of patients with back pain, this subgroup is probably small."

This is an interesting conclusion given that no effort was made to directly examine this hypothesis. Additionally, based on the positive results of their own previous reviews, they felt compelled to make the case that future research is needed to identify this subgroup.[69,124]

Rather than negating the need for future research, the results of this review[138] beg for future studies to match individual patients to interventions with a high probability of success. Given the authors' purpose was to assess only RCT evidence, it is understandable that the initial study that developed the spinal manipulation CPR[12] was not cited. However, the results of this study[12] clearly suggest that this subgroup of patients exists, and that they can be accurately identified

prior to treatment. Based on the detection of an interaction in this validation study, the results of this validation study support the interpretation of Assendelft et al[138] that spinal manipulation is not more effective than other interventions in general. However, the results clearly suggest that a subgroup of patients likely to benefit from this intervention can be accurately identified prior to treatment. Based on the results of the screening examination, and using the presence of at least 4/5 criteria to classify a patient as being positive on the CPR, this subgroup may consist of roughly 10% (47/543) of patients with LBP. However, this is a conservative estimate based on our strict inclusion criteria of a minimum 30% ODQ score at baseline. Presuming the CPR is useful at lower levels of disability, this subgroup may in reality make up 20-30% of patients with LBP, suggesting this subgroup also may not be so small. Although hypothetical, if the studies in this review[138] had considered these characteristics in their inclusion criteria, a positive effect would have likely been detected. Without regard for the classification process, healthcare practitioners, policy makers, and patients who read the results of a review[138] published in a well-respected journal by authors whose interpretation may be viewed as authoritative and final may be falsely misled to believe that spinal manipulation is not effective for any patient with LBP.

6.6 Application of Clinical Prediction Rule to Individual Patients with Low Back Pain

A total of 543 patients were screened for participation in this study, 29% of which (157/543) of which were eligible for participation. This means that 71% of patients (386/543) with LBP were excluded from the study. At first glance, one might question the ability to apply the CPR in a broad spectrum of patients with LBP when 71% of patients were excluded. Clearly, therapists should consider not applying the CPR to patients with LBP who meet one of the exclusion criteria used in this study such as having positive neurologic signs or another red flag that might preclude spinal manipulation as a potential treatment option (i.e. tumor, metabolic diseases, RA,

osteoporosis, prolonged history of steroid use, etc.) However, only roughly 10% of patients with LBP fall into one of these categories (Figure 6). In contrast, 37% of all patients (202/543) who were screened and 52.3% of all ineligible patients (202/386) were excluded for not having at least a baseline ODQ score of 30%. As previously discussed, a higher threshold of baseline disability was incorporated to enable relatively large magnitudes of clinical change to occur, thus minimizing the potential for a floor effect. Given the intent of the CPR to identify patients who experience clinically important changes in disability, application of the CPR clearly becomes less useful at lower levels of disability. For example, although an improvement from 10% on the ODQ at baseline to 5% after treatment represents a 50% improvement, this magnitude of change falls below the MCID of 6 points that has been established for this instrument.[210] Strictly speaking, one could argue that the CPR should not be applied in patients with less than a 30% ODQ score at baseline. To be clear, no data is available to establish a minimum level of disability, below which application of the CPR is no longer useful for decision-making. However, it is likely that the CPR can continue to be useful for some patients with LBP who have less than a baseline ODQ score of 30%. For example, to be conservative, few would argue that an improvement of three times the MCID for the ODQ (i.e. an 18-point improvement) over such a short period of time can be attributed to the favorable natural history of LBP. Based on this consideration, the chance to observe this magnitude of improvement diminishes once patients fall below a baseline ODQ of 20%. Additionally, patients with levels of disability below 20% may more likely represent patients with chronic LBP, who may be more likely to benefit from another intervention such as a stabilization exercise intervention.[225] Based on this rationale, and in light of the overall safety of manipulation in patients who do not have neurologic signs or

another red flag, it is suggested that the CPR is likely still useful for patients with LBP who have at least a baseline ODQ score of 20%.

Perhaps one of the greatest strengths of the spinal manipulation CPR is that it can be applied to an individual patient. All patients who received spinal manipulation were treated with the same manipulative intervention and then classified as being a success or non-success based on whether they achieved at least a 50% improvement in the ODQ, which represents a relatively high threshold of improvement that ensures clinically meaningful change occurred. The accuracy statistics that were calculated can be used for decision-making to identify patients likely to benefit from spinal manipulation. Unlike classic hypothesis testing which involves the comparison of group means using classic inferential statistical procedures such as the t-test and analysis of variance, values for sensitivity, specificity, and positive and negative likelihood ratios are based on the individual patient, thus their interpretation can be readily applied to a single patient. Importantly, the spinal manipulation CPR is the first prediction rule to identify individual patients likely to benefit from this intervention.

6.6.1 Spinal Manipulation Not for All Patients with Low Back Pain

The results of this study do not advocate the use of manipulation as a panacea for patients with LBP, nor do they indicate that manipulation is the only intervention that should be considered for patients who satisfy 3-4 criteria in the CPR. In fact, the findings by Fritz et al[239] suggest there are a cluster of several findings that can be identified a priori identified in the history and physical examination, which if detected in a patient with LBP, suggests that an alternative treatment with a higher probability of success may be warranted. Even patients who meet all five criteria in the CPR will likely need other interventions to complement the use of manipulation to maximize the

patient's outcome. The CPR is intended to identify which patients will receive a large initial benefit from manipulation and does not predict the patient's long-term prognosis. This point can be highlighted based on the history and physical examination findings from Patient #1 in the previously discussed case report,[66] who was suspected to have segmental instability of the lumbar spine based on his history of chronic LBP and positive response to a previous rehabilitation program consisting of spinal stabilization exercises. Some therapists may have chosen not to manipulate this patient based solely on the impression that spinal manipulation would not be helpful, and even perhaps harmful, for a patient with suspected segmental instability. However, consideration of the patient's status with respect to the CPR helped guide the selection of an intervention when, without this information, the therapist may have incorrectly assumed that spinal manipulation should not be considered as a potential treatment option.

In this case, the patient was encouraged to initiate the stabilization exercises again once the acute phase of his LBP was over. Perhaps the manipulative intervention and range of motion exercise in this case served as a catalyst to facilitate his recovery from the acute episode and permit him to initiate the stabilization exercises earlier than he might otherwise have been able. It is highly unlikely this patient's dramatic improvement over such a short period of time could be attributed to the stabilization exercises because these exercises require completion over a longer period of time to demonstrate positive effects.[271]

6.6.2 Clinical Prediction Rule Not the only Criterion to Determine Suitability for Spinal Manipulation

A patient's status with respect to the CPR will in many cases not be the only criterion used to determine which patients clinicians should consider manipulation as a potential intervention. Clearly, other subgroups of patients exist who require interventions with or without the addition of manipulation. For example, a CPR has recently been developed to identify patients likely to benefit from a spinal stabilization approach.[273] It is also important for clinicians to remember the intent of the development of the CPR is to identify patients who are likely to achieve a dramatic improvement in only a very short period of time. For example, the CPR is not designed to identify patients likely to worsen with manipulation. A clinician should not conclude that patients who do not satisfy at least 4/5 criteria in the CPR places the patient at risk for harm or worsening of their status. Clinicians who frequently use manipulation and the ODQ as an outcome measure will attest that improvements on the order of 30-50% over a 1-4 week period of time still represent clinically important change, despite the fact that it does not satisfy the reference criterion for improvement used in the development and validation of the CPR. In these cases, using the CPR alone would fail to identify these patients as potential candidates for manipulation. There may be many scenarios where clinicians appropriately include manipulation in the plan of care, even if the patient does not satisfy the criteria. The CPR, however, will be extremely helpful in assisting clinicians identify patients likely to achieve a dramatic improvement. These are patients therapists will surely not want to miss. Conversely, a therapist may elect not to manipulate a patient who meets 4/5 criteria for a potentially valid reason.

6.6.3 Identifying Patients who may Benefit from an Alternative Intervention

It is interesting to note that only 11.5% (7/61) and 36% (22/61) of patients who received the stabilization exercise intervention alone achieved at least a 50% improvement in the ODQ at the one- and four-week follow-up, respectively. At first glance, this seems to suggest that a stabilization exercise intervention in general may not be as effective as spinal manipulation for patients with LBP. However, treatment in this study was only carried out for four weeks. It has been suggested that the benefits of a stabilization exercise intervention may require completion over a longer period of time to demonstrate improvements.[271] However, a study using a similar stabilization exercise intervention found that only 33% of patients (18/54) demonstrated at least a 50% improvement in the ODQ at the end of eight weeks.[273] It appears that although a stabilization exercise intervention is clearly beneficial for a subgroup of patients with LBP, the effect is generally smaller.

If classification is to be meaningful, perhaps another subgroup of patients with LBP might benefit from a stabilization exercise intervention. Using the same 50% improvement in the ODQ score as the reference criterion, Hicks et al[273] demonstrated that patients likely to benefit from a stabilization exercise intervention tend to 1) have a positive prone instability test, 2) demonstrate aberrant movement during lumbar spine range of motion testing, 3) have an average SLR > 91°, and 4) be < 40 years of age. The positive LR among patients who met at least three of these criteria to identify patients likely to benefit from the stabilization exercise intervention was 4.0 (1.6, 10). With a pre-test probability of 33%, a positive LR of four translates into a post-test probability of 67%, representing a 34% increase in the probability of success when at least three criteria were met. Ideally, a meaningful classification system should be able to distinguish

between somewhat mutually exclusive groups of patients likely to benefit from a specific treatment approach. As the signs and symptoms associated with success from a variety of interventions becomes more clear, future work should examine the mutual exclusivity of patients to various treatment classifications.

6.7 An Alternative Spinal Manipulation Clinical Prediction Rule

Initial development[12] and validation of the spinal manipulation CPR focused on maximizing the positive LR, which is comprised of both sensitivity and specificity values. This resulted in a CPR that maximizes the post-test probability of success. By doing so, equal weight is afforded to making false positive and false negative findings. Table 16 demonstrates that 31 patients achieved at least a 50% improvement in the ODQ at the one-week follow-up, regardless of their status with respect to the CPR. However, 10 of these patients were negative on the CPR (i.e. false negative findings), resulting in a false negative rate of 32.2% (10/31) (Table 16). In other words, using the CPR will cause clinicians to miss 32% of patients who would otherwise benefit from this intervention. This seems to be a high percentage of patients in light of the ease with which the manipulative intervention can be performed and the magnitude of improvement that was missed over such a short period of time.

To minimize the false negative rate, the CPR needs to identify everyone likely to benefit from spinal manipulation, although this will inevitably mean that more patients will receive the intervention but not reach the 50% threshold of improvement, thus increasing the false positive rate. The issue then becomes how to balance the consequence of false negative versus false positive findings. There are very few interventions in a clinician's armamentarium for LBP that can generate change scores on the order of 50% in the ODQ in only one week. Therefore, given

the ease with which the CPR is applied and manipulative intervention can be performed, and in light in the extremely low risks,[124,173,174,176], therapists surely do not want to miss these patients. Therefore, minimizing the false negative rate at the expense of increasing the false positive rate seems reasonable. The MCID for the ODQ has been shown to be 6 points,[210] thus improvements or worsening in status less than 6 points are not considered to be clinically meaningful. A total of 39 patients who received spinal manipulation were classified as a non-success (Table 16), only 10 of which demonstrated higher ODQ scores indicating movement towards increasing disability at the one-week follow-up. Of these 10 patients, only one demonstrated clinically meaningful levels of worsening, and the 6-point increase for this patient just meets the MCID of 6 points for the ODQ.[210] Clearly, patients who receive spinal manipulation are not worsening, regardless of their status with respect to the CPR. Rather they are either failing to improve or achieving clinically meaningful improvements that do not reach the 50% threshold to be classified as a success. In fact, 53.8% (21/39) of patients who received spinal manipulation but not classified as a success achieved clinically meaningful change at the one-week follow-up. Thus it appears that even in the worst case, spinal manipulation is not causing patients to worsen. This finding is also supported in the initial group of patients in which the CPR was developed.[239]

Based on these considerations, a CPR with a sensitivity of 100% needs to be developed to minimize the false negative rate, thus capturing all patients who will achieve at least a 50% improvement in the ODQ at the one-week follow-up. Among the 10 patients who were classified as false negatives (i.e. achieved at least a 50% improvement but did not meet at least 4/5 criteria in the CPR), 8 met at least 3/5 criteria in the CPR. One could simply make a case to lower the threshold to at least 3/5 criteria present to justify an attempt at spinal manipulation; however, this

would still miss two patients who achieved at least a 50% improvement in the ODQ. Both of these patients met 2/5 criteria in the CPR. No patients with either 0 (n=2) or one (n=6) criteria present at baseline achieved at least a 50% improvement in the ODQ (Table 14).

Further analysis of the accuracy of individual items in the CPR illustrates that not all items contribute to the CPR similarly (Table 22, Table 23, Table 24, Table 25, Table 26, Table 27). In fact, among patients with a duration of symptoms less than 16 days who also did not have symptoms distal to the knee, the positive LR for success with spinal manipulation was 12.6 (3.2, 49.8) (Table 38).

Table 38. Accuracy of the CPR to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geq$ 50% improvement on the ODQ. A positive CPR was defined as having a duration of symptoms < 16 days and not having symptoms distal to the knee.

	Success	Non-success	Total (%)
+CPR	20	2	22 (31.4%)
-CPR	11	37	48 (68.6%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: .65 (.47, .79) Sp: .95 (.83, .99)		+LR: 12.6 (3.2, 49.8) -LR: .37 (.23, .61)	

This results in a post-test probability of success of 91%, which is similar to the probability obtained when all five criteria are considered. Interestingly, this suggests that except for ruling out neurologic signs and red flags, a decision can be made to manipulate a patient without ever performing a physical examination! This is not to suggest that the physical examination should be abandoned. Perhaps there is a therapeutic benefit from the examination process itself, and

future study could elucidate a therapeutic role of the physical examination, if any, in patients with LBP. However, the accuracy of these two historical items is quite interesting given the oft-quoted dogma that suggests the decision to utilize spinal manipulation is more complex.

Based on this information, in combination with a goal to create a CPR that was 100% sensitive, an algorithm was developed to minimize the amount of information necessary to influence decision-making, thus maximizing the amount of time saved by clinicians (Figure 16).

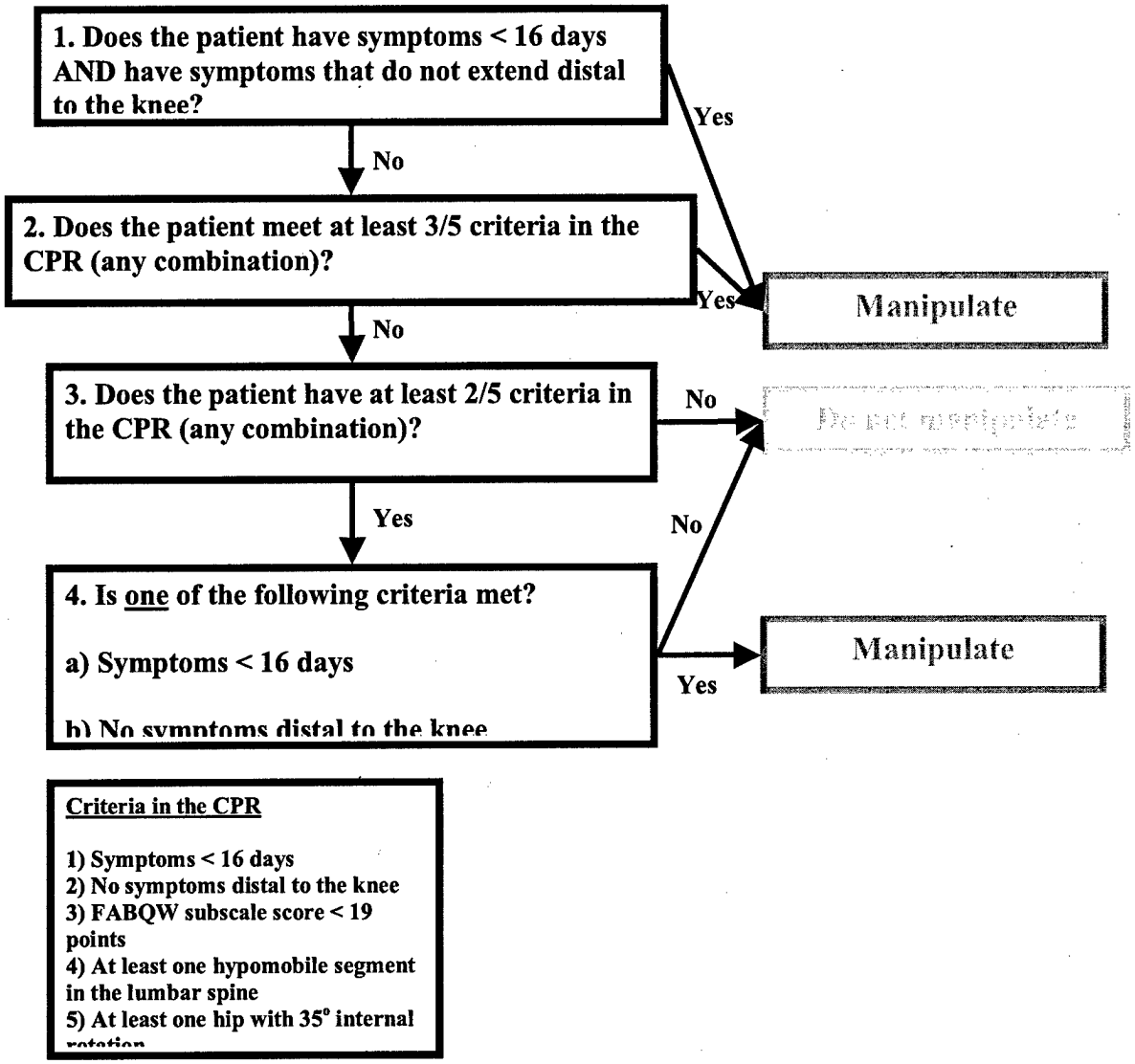


Figure 16. Algorithm to identify all patients likely to benefit from spinal manipulation (i.e. 100% sensitive).

The two-by-two contingency table generated from the patients in this study had this algorithm been used for decision-making is depicted in Table 39.

Table 39. Accuracy of the CPR algorithm to identify patients likely to benefit from spinal manipulation at the one-week follow-up. Success was defined as $\geqslant$ 50% improvement on the ODQ. A positive CPR was defined as satisfying a decision point in the algorithm that would result in a recommendation to use spinal manipulation.

	Success	Non-success	Total (%)
+CPR	31	25	56(80.0%)
-CPR	0	14	14 (20.0%)
Total (%)	31 (44.3%)	39 (55.7%)	70
Sn: 1.0 (.89, 1.0) Sp: .36 (.23, .52)		+LR: n/a -LR: n/a	

(Return to p. 159, 159)

Had this algorithm been used, 80% of patients (56/70) would have received spinal manipulation, decreasing the false negative rate to 0% (Table 39). However, the false positive rate increases from 5.1% (2/39) using the 4/5 threshold (Table 16) to 64.1% (25/39) with the algorithm (Table 39). In essence, therapists would have manipulated an additional 23 patients who would not have achieved at least a 50% improvement in their ODQ to capture an additional 10 patients who would have responded to spinal manipulation.

Development of an alternative CPR does not undermine the usefulness of the previously established 4/5 as a valid cut-off to establish an overall positive test. Rather they have different interpretations. Following the alternative algorithm allows clinicians to be assured they are

identifying all patients likely to benefit from spinal manipulation. Obviously, therapists may still not elect to use this intervention, and patients do not have to provide consent. Rather, based on the results of the algorithm, clinicians should be expected to offer this management approach to their patients when the decision points are met. The therapist and patient can then decide the most appropriate course of action through shared decision-making.

Using the 4/5 cut-off to determine when to recommend spinal manipulation may have the most value for clinicians who continue to be reluctant to routinely use spinal manipulation in clinical practice or patients who are unsure as to whether they should receive this intervention. Clinicians can confidently say with 91% certainty that patients who meet at least 4/5 criteria in the CPR will achieve at least a 50% improvement in their ODQ by the end of one week. This has important implications for decision-making and to establish a patient's prognosis. It can be argued that clinicians who do not offer spinal manipulation for these patients are withholding an intervention that has a high probability of being effective. Identifying patients who meet at least 4/5 criteria in the CPR may help persuade reluctant clinicians to provide spinal manipulation for these patients. Failing to do so will forgo a practical guarantee that the patient would achieve at least a 50% improvement in the ODQ by the end of one week, a response that cannot likely be achieved with an alternative stabilization exercise intervention.

6.8 The Consequences of Misperceptions Regarding Spinal Manipulation

The Federation of State Board of Physical Therapy (FSBPT) has the responsibility to develop and maintain the National Physical Therapist Examination (NPTE), which is the required examination for all physical therapists seeking licensure in the United States. The purpose of the NPTE is to "assess basic entry-level competence of the licensure candidate who has graduated

from an accredited program of physical therapy".[274] The FSBPT's role is to provide examination services to state regulatory boards charged with regulating physical therapy practice and to provide a standardized mechanism to insure a comparable standard for competence across jurisdictions. Unfortunately, misperceptions regarding the use of spinal manipulation have profound consequences on the inclusion of content related to manipulation on the NPTE. To better understand these implications, it is helpful to illustrate the process by which the NPTE is developed and constructed.

6.8.1 Development and Construction of the National Physical Therapist Examination

6.8.1.1 Job Analysis Survey

The NPTE is developed by physical therapists who serve on a variety of FSBPT committees. Every few years, a group of physical therapist subject matter experts (SME) is assembled to determine the knowledge, skills and tasks consistent with physical therapy scope of practice.[275] *The Guide to Physical Therapist Practice*[68] is used as the primary resource to organize and describe activities. A final summary list is formed using a Delphi consensus approach. After being pilot tested among a representative sample of therapists throughout the United States, a job analysis survey is distributed to a random sample of licensed physical therapists.[275] Therapists are asked to rank activities according to three criteria: 1) acquisition, 2) criticality, and 3) frequency. Acquisition is concerned with identifying whether the knowledge requirements and skills necessary to perform a particular activity are acquired during entry-level education or represent an advanced skill acquired at some point during clinical practice. Criticality represents the extent to which incorrect performance of an activity could cause the patient psychological or physical harm. Finally, frequency identifies how often a particular activity is performed.

6.8.1.2 Development of Content Outline

The scores for each criterion are combined in an overall composite score for each activity based on a weighting factor assigned to each criterion. The criterion related to acquisition is weighted most heavily. The activities are then rank-ordered based on the composite score, and a cut-off is established to identify the activities to be included on the content outline.[275] Activities that do not exceed the threshold score are judged to be inconsistent with entry-level practice and omitted from the content outline. The content outline serves as blueprint for construction of the NPTE, listing specific content areas and the number of questions for each area that must be included on the examination.[275]

Although spinal manipulation is underutilized among physical therapists in general,[165-167] this phenomenon is particularly pronounced among entry-level therapists.[181] Data from the most recent *Job Analysis for U.S. Physical Therapy Practice*[181] completed in 2002 demonstrates that, on average, therapists perceive spinal manipulation to be an advanced skill to be learned through post-professional education and that incorrect performance of these interventions will cause "severe psychological or physical harm".[181] Only 11.7% of therapists with 1-2 years of experience report using spinal manipulation on a daily or weekly basis, and 62.8% of therapists with this level of experience do not utilize these skills at all.[181] Unfortunately, utilization rates do not substantially improve among therapists with greater than two years of experience (25.3% and 50.9%, respectively).[181] Based on its low utilization rate, high level of perceived harm, and perception that these skills are not consistent with entry-level practice, spinal manipulation was omitted from the content outline currently used to construct the NPTE.[276] Manipulation of the extremities is excluded from the current content outline for similar reasons.[181] It is curious that

some standards seem to have been arbitrarily applied to manipulation, but not other practice areas. For example, questions related to pediatric physical therapy are rightly included on the content outline; however, only a small percentage of physical therapists work in this practice setting, thus utilization rates across physical therapy practice as a whole are not high.

It is unclear if the FSBPT considered the prevalence of manipulation among subgroups of physical therapists. For example, it would be expected that manipulation should be more prevalent among therapists in a predominantly outpatient orthopaedic practice setting than therapists in a primarily neurologic or pediatric setting, for example. If subgroups of therapists were not considered, utilization rates of manipulation may have been artificially low. Therefore, conclusions drawn from the job analysis survey regarding decisions as to which content should be included on the content outline may have been incorrectly made. Although "manual therapy" is specifically listed, the definition is limited to "techniques including spinal and peripheral mobilization, manual traction, and techniques of soft tissue mobilization."[276] Interpretation of the term "mobilization" has apparently excluded the concept of a Grade V mobilization according to the Maitland classification,[186] thus no items specific to "high-velocity thrust" techniques are permitted for either the spine or extremities.

6.8.1.3 Development of Test Items

After the content outline is developed, members of the Item Writing Review Committee representing a broad range of practice settings develops items based on the content areas and their associated weights. Items are reviewed by a Regional Coordinator, after which approved items are reviewed by the Item Bank Review Committee. After necessary revisions are made through the various levels of individual and committee oversight, a subset of items is eventually

included in the item bank. These items are then grouped into pretest blocks where the items are included on the examination but do not contribute to a candidate's score. Candidates are unaware which items are included in the pretest block. Statistical properties of the items are determined, and appropriate revisions are made to maximize the ability of the item to distinguish between successful and non-successful candidates. Revised items are then re-tested in a subsequent examination administration. Once the statistical requirements for an item are deemed acceptable, it can be included as a testable item that contributes to the candidate's score. Approximately 60-70% of pre-test items eventually appear as a testable item on the examination.[275] The Examination Construction and Review Committee then uses the content outline to make the determination as to which testable items will appear on a given examination.

6.8.2 The "Evidence Gap"

The FSBPT's aim is to involve a "large, representative group of practicing physical therapists and physical therapist assistants and other professionals at every stage of examination development [to] ensure that the examinations are relevant to the current practice of physical therapy."[274] However, it is important to make a distinction between the examination reflecting "current practice" versus "best practice". The examination can only reflect current evidence presuming practice patterns among are consistent with the evidence. It seems logical that the presumed intent and assumption of the FSBPT is that "current practice" ultimately will reflect practice patterns consistent with current evidence in the literature. However, this does not appear to be the case. In fact, the "evidence gap" between "current practice" and "best practice" currently appears to be wide.

The following physical agents are specifically listed on the current content outline implemented in 2002,[276] which insure that specific items related to these agents will be included on the examination: 1) intermittent compression, 2) superficial thermotherapy (e.g., hot packs, paraffin, and Cryotherapy), 3) ultrasound including phonophoresis, 4) electrical stimulation including iontophoresis, 5) biofeedback, 6) mechanical modalities (e.g. traction, tilt table/standing frames, continuous passive motion), and 7) whirlpool/Hubbard tank. There is growing evidence for the effectiveness for electrical stimulation to improve muscle function for a variety of neuromusculoskeletal disorders.[277-282] However, the common theme throughout the recently published Philadelphia Panel evidence-based guidelines on selected interventions for low back,[283] neck,[284] shoulder,[285] and knee[286] was that there is almost no evidence for the use of physical agents in the management of musculoskeletal disorders, yet many of these agents continue to be listed in the content outline.[276] Hurwitz et al[287] found no additional benefit for a variety of physical agents in the management of LBP compared to the use of spinal manipulation alone. In fact, interventions such as ultrasound have been studied at length for a variety of musculoskeletal disorders and found to be ineffective.[288] Until evidence was recently published demonstrating some short-term effectiveness for superficial heat in patients with acute LBP,[289,290] no evidence existed for its effectiveness.[283-286] Although there is limited evidence to support the use of biofeedback in the management of urinary incontinence,[291-293] recent evidence from RCTs[294,295] and a systematic review[296] have questioned the usefulness of this modality. Except for patients with knee osteoarthritis,[286,297] the use of TENS to improve pain for a variety of musculoskeletal disorders has been found to be ineffective.[298,299] Furthermore, there is also conclusive evidence that continuous passive motion is of little value after total knee arthroplasty.[300-304] Manual therapy interventions such as joint and soft tissue mobilization are

specifically included in the content outline,[276] yet only limited evidence exists to support their effectiveness. For example, the use of massage for a variety of musculoskeletal disorders appears to be limited,[305,306] and joint mobilization seems to be less effective than manipulation for patients with LBP.[73,80] In contrast, clinical practice guidelines in the United States,[21,22,142] New Zealand,[22,143] and the United Kingdom[146] all recommend spinal manipulation for patients with non-radicular acute LBP based on a systematic assessment of the evidence. However, this content is omitted from the current content outline for the NPTE.[276] Despite this paradox, non evidence-based interventions will continue to be included as long therapists report using these skills in the job analysis survey.

One particular requirement by the FSBPT in the development of the NPTE further exacerbates the "evidence gap". Item writers are specifically instructed that only textbooks considered to be "authoritative" may be used to reference items on the NPTE. Although no precise interpretation of the term "authoritative" could be found, the standard appears to require that it be widely used and accepted in physical therapy practice. Original journal articles from the peer-reviewed literature are specifically excluded as a potential source from which items can be developed. This seems odd given the emphasis on evidence-based practice. Clinical expertise communicated in non-peer reviewed literature such as textbooks is the lowest level of evidence in the evidence hierarchy. Because there is no peer-review requirement, textbooks often strongly reflect an author's own personal bias rather than current evidence and may be inundated with authoritarian approaches that entirely lack evidence for their use. Additionally, evidence from an original manuscript in the peer-reviewed literature may be 1-2 years old by the time a manuscript has been written, submitted for publication, navigated its way through the peer-review process, and

finally published in a journal for widespread dissemination to clinicians. On the other hand, assuming an author of a textbook actually makes an effort to reflect evidence, they will often lag behind the evidence by 5-10 years because of the time required to assemble the volume of information from contributing authors and complete the editing and review process by the publishing company. This gap is further widened when editions are infrequenlt updated. There is no higher standard for authority than a manuscript that has successfully navigated its way through the peer-review process. Although some aspects of physical therapy practice remain constant across time, evidence for physical therapy practice is growing at an ever increasing rate. Excluding orginal articles from the peer-reviewed literature only magnifies the "evidence gap". As a minimum, the FSBPT should encourage item writers to use published manuscripts from the peer-reviewed literature.

6.8.3 The Case for Spinal Manipulation as an Entry-level Skill

The false notion that spinal manipulation poses undue risks to patients and that these skills are perceived to require advanced training[181] likely accounts for the reason why clinical practice continues to lag behind mounting evidence that suggests spinal manipulation should be widely used. This "evidence gap" may largely be attributed to the failure of entry-level programs to teach manipulation.[307-310]

6.8.3.1 Prevalence of Spinal Manipulation in Entry-level Curricula

In 1970, Stephens[307] surveyed physical therapy programs and found that only 9 out of 51 programs (17.5%) offered instruction in manipulation, however the operational definition of manipulation was not clearly defined in this study. Ben-Sorek and Davis[308] conducted a similar survey in 1986 and found that 93% of programs incorporated some instruction in "mobilization",

defined as a "skilled passive movement to a joint". Seventy-two percent of respondents expressed an interest in expanding the amount of instruction in joint mobilization.[308] More recently, Bryan et al[309] determined that 103 out of 104 physical therapy programs taught "spinal mobilization", which was defined as "an act of imparting movement, actively or passively to the joints or soft tissues of the spinal column". However, 96% of the faculty reported having received their spinal mobilization training through continuing education.[309]

6.8.3.2 Is the Glass Half-Empty or Half-Full?

Boissionault et al[310] recently completed a follow-up study to specifically assess the extent to which manipulation (i.e. high-velocity thrust techniques) was included in entry-level curricula. Unfortunately, only 44% of programs still report teaching these skills.[310] Forty-two percent of programs did not respond to the survey.[310] Although purely speculative, one might suspect that some programs not teaching manipulation may perhaps have been more reluctant to respond, thus the true proportion of programs teaching these skills could actually be lower.

Faculty reasons for not including manipulation in their curricula include a belief that it is not an entry-level skill (45%), lack of time (26%), lack of qualified faculty (7%), and perceived lack of scientific evidence regarding efficacy (7%).[310] However, the notion that spinal manipulation should be considered an advanced skill only to be learned through post-professional continuing education courses, fellowships/residency programs, or post-professional degree programs is clearly unwarranted. Reviews of the evidence suggest that the safety and effectiveness of spinal manipulation is not dependent on the type of practitioner, technique used, or years of experience,[21,184,311] and patients likely to benefit from spinal manipulation can be identified prior to treatment with increasing certainty.[12] Spinal manipulation is not the exclusive domain of any

single profession, nor is it an esoteric skill that requires years of training to develop. Rather, it is a motor skill that most entry-level therapists should be able to acquire with adequate practice. Fortunately, the incorporation of spinal manipulation into entry-level curricula appears to be on the rise. Among programs that do not presently include manipulation in their curricula, 51% report plans to do so in the near future.[310]

6.8.3.3 The Evaluative Criteria and Spinal Manipulation

The Commission on Accreditation of Physical Therapy Education (CAPTE) has recently made an effort to get more entry-level physical therapy programs to incorporate manual therapy training, including spinal manipulation, into their entry-level therapist curriculum.[312] The American Physical Therapy Association (APTA) and the American Academy of Orthopaedic Manual Physical Therapists (AAOMPT) also formed a manipulation task force in 1998 to increase and enhance the level of instruction in spinal manipulation in physical therapy education. However, the Evaluative Criteria[313] currently contains the more broad term of "manual therapy" (3.8.3.28, f.). The CAPTE has interpreted this criterion to mean that programs can satisfy the current language by using lower velocity, joint mobilization or soft tissue techniques, thus are not required to include any high-velocity thrust techniques. This seems unusual in light of *The Guide to Physical Therapist Practice*,[68] which clearly defines mobilization/manipulation to include "a small-amplitude/high-velocity therapeutic movement,"[68] and evidence that suggests joint mobilization is less effective than manipulation for patients with LBP.[73,80] Clearly, this discrepancy should be corrected in the next update of the evaluative criteria to require the teaching of at least a core set of high-velocity thrust techniques for spinal disorders.

6.8.3.4 The Vicious Cycle: Competing Demands for Curricular Attention

Entry-level physical therapy programs only have a finite number of hours in their curricula. Many content areas compete for increased emphasis, and faculty continuously negotiate the amount of time allocated to any single content area to insure the curriculum as a whole meets the CAPTE Evaluative Criteria[313] and the program's unique mission and needs. For sure, one of the key objectives of the curriculum is to prepare students for success on the licensure examination and competent physical therapy practice. However, the need to prepare students for success on the examination seems to unintentionally contribute to a vicious cycle that further discourages the inclusion of manipulation in entry-level curricula. When entry-level programs fail to teach these skills, it seems logical that recent graduates may not feel competent using these skills in clinical practice. Underutilization of these skills is then reflected in the job analysis survey, resulting in the omission of this content from the examination. During the negotiation process to allocate time to each content area in a program's curriculum, a case can be made that because content related to manipulation is not tested on the NPTE, it does not deserve attention in the curriculum. At the same time, however, other non evidence-based interventions currently included on the NPTE continue to be emphasized in the curriculum to prepare students for these items on the examination. It seems logical that at least in the short-run, graduates are likely to practice in a manner consistent with their educational experience. Thus utilization of a number of non evidence-based interventions continues be reflected in the job analysis survey, and the vicious cycle repeats itself. It seems this pattern is likely to continue unless modifications in the development process of the NPTE are made.

6.8.4 The Future of Spinal Manipulation and the National Physical Therapist Examination

Using a retrospective, feedback-based approach based on a job analysis survey may not be the ideal mechanism to achieve an examination that reflects "current practice". Because the job analysis survey is only conducted every 5-6 years, the opportunity to reflect changes in practice patterns is quite infrequent. Furthermore, it takes approximately 3-4 years from the time a job analysis survey is completed until an item navigates its way through the review process for consideration as a testable item on the NPTE. Therefore, even presuming spinal manipulation is included on the next content outline presumably sometime around 2008, testable items related to this intervention will not appear on the NPTE for another 8-10 years from now. If the next job analysis does not reflect increased utilization of these skills and a more realistic perception of the risks, 16-20 more years will pass before content related to spinal manipulation could be considered for inclusion.[275]

Rather than establishing a level of competence that might meet the expectations of the patients we serve, items on the NPTE are intentionally written only to distinguish between therapists who are "minimally competent" to practice physical therapy and those who are not. Item writers are constantly encouraged to identify in their own mind the characteristics of a "minimally competent" therapist, evoking images of someone who may not cause egregious harm to a patient, but on the other hand may not have the requisite skills to actually benefit the patient. Undoubtedly, the minimally competent therapist would not be someone that most therapists would like to have as a colleague or would want as their therapist. However, item writers are continually reminded to consider the "minimally competent" therapist when writing items to

minimize having items that assess content beyond entry-level practice. However, given the recent proliferation of evidence for physical therapy practice, it seems odd that we have such low expectations of what entry-level practice should be. It is understandable that we do not expect entry-level therapists to be clinical experts and be tested using a standard that might be used on a board-certification examination, for example. However, surely a more reasonable level of competence could be used. Perhaps the standard should represent a therapist who most therapists would not object to having as a colleague, or who would agree that the therapist could competently practice physical therapy for a majority of their patients. It is unlikely that our profession and the public at large is best served using the "minimally competent" therapist as the standard by which our profession licenses therapists.

6.8.5 Maintaining the Status Quo Not an Option

The exclusion of manipulation content on the NPTE[276] and its de-emphasis in entry-level curricula[310] all provide legislative and political fodder for opposition groups who wish to limit physical therapy scope of practice and deny therapists the right to use manipulative interventions in clinical practice. Additional efforts must be made to encourage clinical practice being consistent with the evidence. The failure to do so will only jeopardize our profession's rapid transition toward becoming a doctoring profession with a vision to provide direct access physical therapy services.

6.8.5.1 Evidence-based Practice: The Ideal Minimum Standard of Competence

Perhaps a feed-forward mechanism in which a representative panel of subject matter experts in various practice areas is assembled would be more helpful to achieve the FSBPT's purpose and protect the public safety. Rather than focusing on "current practice" based on the job analysis

survey, this panel would develop a content outline that includes tasks and roles consistent with "best practice" according to the currently available evidence. Working alongside the CAPTE, this process would indirectly facilitate evidence-based practice as the minimum standard of competence, rather than settling for a minimum standard designed primarily to protect the public from the most egregious deficiencies in competence. Entry-level programs would indirectly then have to insure their curricula remained consistent with the evidence to prepare their graduates for a more current examination. Content included on the NPTE will almost certainly continue to lag behind the evidence by at least 8-10 years as long as the current approach remains in place.

6.9 Incorporating the Spinal Manipulation Clinical Prediction Rule into Clinical Practice

Clinicians need to be proficient in the manipulative intervention and familiar with the individual items and overall decision-making process involved in the application of the CPR. Because the results of this study provide clinicians with a practical and evidence-based approach to quickly identify the subgroup of patients with LBP likely to likely benefit from this intervention prior to treatment, combined with the ease in which the manipulative intervention can be performed, incorporation of the CPR into clinical practice should be a reasonable task. However, despite the high level of evidence that exists for its use when decision-making is based on the CPR, some therapists will inevitably be reluctant to utilize even this single manipulative intervention. However, the results of this study directly address the potential reasons why therapists may be reluctant to utilize this intervention in the management of patients with LBP.

6.9.1 Risk of Worsening with Spinal Manipulation

The perception among clinicians is that the risks associated with spinal manipulation are greater than those associated with alternative interventions such as exercise.[181] Evidence suggests this

perception likely contributes to therapists being reluctant to incorporate these skills in clinical practice.[181] However, the evidence clearly does not support the notion that spinal manipulation poses an unreasonable risk to patients with non-radicular LBP. The risk of a serious complication from spinal manipulation such as cauda equina syndrome is extremely low,[124,173,174,176] estimated at approximately 1 per 100 million manipulations.[124,174] In contrast, the risk of sudden death from exercise has been established to be roughly 1 per 1.5 million episodes of physical exertion, an almost 10-fold increase.[314] Not only is the risk of cauda equina syndrome substantially lower than that associated with sudden death from exercise, but few would argue that a surgically correctable cauda equina syndrome is a more serious adverse outcome than the irreversible condition of death. Perhaps this perception is perpetuated in many professional continuing education courses that teach spinal manipulation and entry-level programs that teach these skills. The notion that spinal manipulation is "dangerous" and should only be practiced by practitioners with "advanced" training may serve to heighten the "expert" clinician's ego; however, it offers little to the patient who would otherwise benefit from a potentially effective intervention.

6.9.1.1 Clinical Factors Associated with a Failure to Improve with Spinal Manipulation

Given the extremely low risks of a serious complication from spinal manipulation, it has been suggested that the greatest risk may be a worsening in the patient's status, or simply the failure to improve.[239] Studies reporting the prevalence of adverse effects of spinal manipulation have not described the clinical presentation of patients whose status was worsened as a result of this intervention. If the clinical presentation of patients unlikely to benefit from lumbar spine manipulation could be characterized, this information would be particularly useful for clinicians who are reluctant to utilize this intervention. Among the same group of patients in which the spinal manipulation CPR was developed,[12] Fritz et al[239] conducted a study to identify factors

from the history and physical examination that were associated with a worsening, or lack of improvement in the clinical status of patients with LBP who were treated with spinal manipulation. Seventy-one patients with non-radicular LBP (mean age 37.6 ±10.6 years, mean duration of symptoms 41.7 ± 54.7 days) received a standardized baseline assessment including history, self-reports, lumbar and hip range of motion, and various diagnostic tests to assess dysfunction in the lumbopelvic region. All patients were treated with spinal manipulation for a maximum of two sessions. Patients who did not show greater than five points of improvement in the ODQ were considered to have failed to improve with the manipulative intervention. Baseline variables were tested for significant univariate relationship with the outcome of the manipulation. Variables showing univariate significance were entered into a logistic regression equation and adjusted odds ratios were calculated to determine the explained variability in outcome with these 6 factors.

Only 28% (20 patients) failed to improve with manipulation, thus 72% showed meaningful clinical improvement after manipulation. These results demonstrate that the majority of patients with LBP seem to improve with manipulation, even if patients are manipulated without regard to the history and physical examination findings. Importantly, no patients in this study experienced any serious adverse effects of the manipulation and only two patients worsened when using a greater than five-point increase in the ODQ score as the criteria to define worsening,[210] providing additional evidence that manipulation appears to be a relatively safe intervention in patients with LBP. However, 6 variables were identified as being significantly related to failure to improve with manipulation and thus predictive of outcome: 1) longer symptom duration, 2) having symptoms in the buttock/leg, 3) not having lumbar hypomobility, 4) less hip rotation range of

motion, 5) less discrepancy in left-to-right hip internal rotation range of motion, and 6) a negative Gaenslen's sign.[239] Interestingly, only one diagnostic test for the lumbopelvic region, the Gaenslen sign, was associated with outcome, providing additional evidence of the futility of these tests. The resulting logistic regression model that incorporated these findings explained 63% of the variability in manipulation outcome.

From the history, the most important factors associated with failure to improve with manipulation were a longer duration of symptoms, and the presence of symptoms distal to the low back. The improved effectiveness of manipulation in patients with more acute symptoms has been identified in subgroup analyses of previously published RCTs.[80,84] Spinal manipulation is believed by some to be contraindicated for patients with sciatica.[21,315] Patients with signs of nerve root compression were excluded from this study; however, patients with symptoms into the buttock or leg(s) were more likely to fail to improve with manipulation.[239] Ninety-percent of patients who failed to improve had symptoms distal to the low back, and 40% had symptoms distal to the knee, compared with 61% and 20%, respectively, for patients who improved.

Similar to the factors in the CPR to predict success with manipulation,[12] relatively few physical examination findings were significantly associated with a failure to improve. Most of the physical examination findings associated with treatment failure were related to the presence of less hip internal and external rotation range of motion and less discrepancy in internal rotation range of motion between the left and right hips. Although several investigators have suggested a link between limited hip rotation range of motion and the presence of LBP,[236-238,240,241,316-318] additional research is needed to explore the relationship between the range of hip rotation and

outcome from manipulation; however, these data[239] suggest that patients with a characteristic pattern of less hip internal and external rotation range of motion and less discrepancy in internal rotation range of motion between the left and right hips may be more likely to respond to an intervention other than spinal manipulation. The important result of this study is that clinicians can use these preliminary findings to a priori identify those patients who may be more likely to benefit from an intervention other than spinal manipulation, information which can assist clinicians in decision-making.

6.9.1.2 Quantifying the Risk of Worsening from Spinal Manipulation

Characterizing the factors associated with a failure to improve with spinal manipulation is helpful for decision-making and may serve to dampen the false notion that utilizing this intervention poses unnecessary risks to patients with LBP. Although the risk is extremely low,[124,173,174,176] the "seriousness" of the albeit almost negligible risk among patients with non-radicular LBP may contribute to the reluctance among some clinicians to routinely utilize these skills.[181] Therefore, it would be helpful to quantify this risk compared to an alternative intervention believed to be "less risky" such as a stabilization exercise approach.

To characterize the risks associated with spinal manipulation, researchers and clinicians have historically relied on a rather defensive position by defining the risk in terms of experiencing a serious complication. However, clinicians who routinely use these skills will readily attest to the notion that spinal manipulation is safe and effective. But more importantly, because many patients seem to experience a somewhat dramatic improvement, perhaps the failure to offer this intervention may actually place the patient at risk for not achieving an optimal outcome. Thus it

would be interesting to consider a more "offensive", diametrically opposed way of thinking, and actually quantify the risk of NOT of not offering spinal manipulation to patients with LBP.

To characterize the risk of failing to offer spinal manipulation for patients with LBP, the MCID of 6 points for the ODQ[210] was used to classify patients in both groups as to whether they improved, worsened, or remained unchanged in their clinical status at the one- and four-week follow-up examination (Table 40).

Table 40. Number (percent) of patients in each group who improved, worsened, or remained unchanged in their clinical status at the one- and four-week follow-up. Improvement and worsening was defined as changes $\geqslant$ 6 points and $\leqslant$ 6 points in the ODQ, respectively. Otherwise, patients were classified as unchanged.

	One-week			**Four-week**		
	Improved[1]	**No change**	**Worsened[2]**	**Improved[3]**	**No change**	**Worsened[4]**
Manipulation Group (n=70)	52 (74.3%)	17 (24.4%)	1 (1.4%)	57 (81.4%)	11 (15.7%)	2 (2.9%)
Exercise Group (n=61)	31 (50.8%)	23 (37.7)	7 (11.5%)	37 (60.7%)	17 (27.9%)	7 (11.5%)

[1] $\chi 2$=7.7 (p=.005), odds ratio=2.8 (1.3, 5.8) (p=.006)
[2] $\chi 2$=5.7 (p=.017), odds ratio=8.9 (1.1, 74.9) (p=.043)
[3] $\chi 2$=6.9 (p=.008), odds ratio=2.8 (1.3, 6.3) (p=.01)
[4] $\chi 2$=3.8 (p=.052), odds ratio=4.4 (.88, 22.1) (p=.071)

(Return to p. 178, 178, 179, 179, 179, 179, 180)

Only 25.7% (18/70) of patients failed to demonstrate clinically meaningful improvement with spinal manipulation at the one-week follow-up, thus 74.3% (52/70) demonstrated improvement (Table 40). In contrast, 49.2% (30/61) of patients who received the stabilization exercise intervention alone failed to improve at the one-week follow-up (Table 40). Similar figures are

observed at the four-week follow-up (Table 40). These results demonstrate that even if clinicians randomly manipulate patients with LBP, paying no attention to the history and physical examination, patients are likely to demonstrate clinically meaningful improvements. No patients who received spinal manipulation experienced a serious adverse event, and only one patient experienced a worsening in status, and this patient just met the threshold by experiencing a 6-point increase in ODQ at the one-week follow-up (Table 40). 12% of patients who received the stabilization exercise intervention experienced a worsening in status at the one-week follow-up compared to only 1% of patients who received spinal manipulation (p=.017). A similar trend was observed at the four-week follow-up (p=.052). Alternatively, 74% of patients who received spinal manipulation experienced clinically meaningful levels of improvement at the one-week follow-up compared to only 51% who received the stabilization exercise intervention (p=.005). A similar finding is noted at the four-week follow-up (p=.008) (Table 40).

Perhaps the best way to illustrate the risks associated with spinal manipulation is to determine the odds of experiencing a worsening in status based on whether the patient received spinal manipulation or the stabilization exercise intervention alone. In this case, the risk factor was defined as not receiving spinal manipulation, and the "adverse" outcome was defined as experiencing at least a 6-point worsening on the ODQ at the one-week follow-up. The odds ratio associated with a worsening in status at the one-week follow-up for patients not receiving spinal manipulation was 8.9 (1.1, 74.9) (p=.043) (Table 40). This means that patients who received the stabilization exercise intervention alone were almost 9 times as likely to experience a worsening in status compared to patients who received spinal manipulation. Because there is little theoretical rationale for why the stabilization exercise intervention is actually harmful, a more

accurate interpretation might be to say that the failure to offer spinal manipulation to patients with LBP in general places patients at a 9-fold increased risk of worsening. A similar trend is observed at the four-week follow-up. Alternatively, patients who received spinal manipulation were almost three times as likely to experience clinically meaningful levels of improvement at the one- (p=.006) and four-week (p=.01) follow-up examination (Table 40).

Another approach to characterize the risk of failing to offer spinal manipulation for patients with LBP is to characterize the odds of success with spinal manipulation compared to an alternative intervention such as a stabilization exercise program. The number of patients in each group who were classified as a success at the one-and four-week follow-up is depicted in Table 41.

Table 41. Number (percent) of patients in each group who were classified as a success at the one- and four-week follow-up. Success was defined as ≥ 50% improvement in the ODQ score.

	One week*		**Four weeks****	
	Success[1]	**Non-success**	**Success[2]**	**Non-success**
Manipulation Group (n=70)	31 (44.3%)	39 (55.7%)	44 (62.9%)	26 (37.1%)
Exercise Group (n=61)	7 (11.5%)	54 (88.5%)	22 (36.1%)	39 (63.9%)

[1] $\chi 2$=17.0 (p<.001), odds ratio=6.1 (2.4, 15.4) (p<.001)
[2] $\chi 2$=9.4 (p=.002), odds ratio=3.0 (1.5, 6.1) (p=.003)

(Return to p. 181, 181)

44.3% (31/70) experienced at least a 50% improvement in the ODQ at the one-week follow-up compared to only 11.5% (7/61) of patients who received the stabilization exercise intervention

(p<.001). By the four-week follow-up, 62.9% (44/70) of patients who received spinal manipulation had a successful outcome comapred to only 36.1% (22/61) of patients who received the stabilization exercise intervention (Table 41). The corresponding odds ratios demonstrate that patients who received spinal manipulation were 6 times as likely to experience a successful outcome at the one-week follow-up (p<.001) and three times as likely at the four-week follow-up (p=.003) (Table 41).

Interestingly, these data pay no attention to decision-making related to the CPR. It seems logical that odds of changes in clinical status might be further magnified when decisision-making related to the CPR is considered (i.e. considering only patients classified as positive on the CPR). No patients classified as positive on the CPR experienced a worsening in status, thus calculation of the odds ratio among only patients classified as positive on the CPR is indeterminable. However, based on these data, clinicians can be virtually certain that the decision to use spinal manipulation in these patients will not lead to a worsening in status. The odds of improvement among only patients classified as positive on the CPR increases to 22 (2.5, 190.4), and the odds of success increases to 73.5 (11.1, 485.9). This means that patients classififed as positive on the CPR are 74 times as likely to experience a successful outcome if they receive spinal manipulation than if they receive a stabilization exercise intervention.

A historically defensive position has been used to characterize the risks associated with spinal manipulation; however, the results of this study suggest that the risk of failing to routinely offer this intervention for patients with LBP is real and that a more offensive approach is warranted to describe these risks. Not only does spinal manipulation not expose patients to unnecessary risks

of serious complications, but the failure to widely utilize this intervention in patients with LBP actually increases a patient's risk of worsening. Alternatively, using this intervention significantly increases the odds of experiencing clinically meaningful levels of improvement and a successful outcome.

6.9.1.3 Spinal Manipulation and the Informed Consent Process

Despite evidence that suggests spinal manipulation is beneficial among patients classified as positive on the CPR, and that the risks are extremely low,[124,173,174,176] the perception of harm must be considered. Clearly, the benefit of spinal manipulation among patients who meet at least 4/5 criteria appear to outweigh the very small risk. However, as with any intervention, the patient should be informed of the risks and benefits to make an informed decision. It is important to let the patient know that according to the current understanding of the problem, serious injury is extremely rare. Therapists are cautioned against overstating the risks and unnecessarily heightening the patient's level of anxiety. It may help therapists put things into perspective by considering the risk/benefit ratio of other commonly prescribed treatments, such as NSAIDS. Clinicians could say something like, "I would like to proceed with a manipulative intervention designed to increase motion and decrease pain in your low back. The risk of this procedure is extremely low. In fact, the risk of having a serious adverse side effect from taking NSAIDs is greater than the risks associated with manipulation. If you are uncomfortable in anyway please let me know. Furthermore, based on your clinical examination, you have some factors that suggest manipulation is likely to be very helpful to improve your pain and function in only a few days." It is also important to document that the patient consented to manipulative intervention procedures and that any screening tests, if performed, were negative. Therapists should not view informed consent as a "line in the sand", so to speak, after which clinicians are free to do

whatever they wish. Informed consent is really an ongoing process of communication between the clinician and patient. The demeanor and goals of the patient, nature of referral, skill of the therapist, and bias of the referring provider must all be weighed in the context of the overall decision-making process in the determination to utilize mobilization/manipulative interventions.

6.9.2 Are the Benefits of Spinal Manipulation Worth the Effort?

The results of this study clearly demonstrate improvements in outcome when a patient's status with respect to the CPR is considered. However, determination of whether the benefits are worth the effort required to use the CPR in clinical practice is also an important consideration.[233] The number needed to treat (NNT) statistic is a useful statistic to make this determination. The NNT represents the number of patients a clinician must treat with the intervention of interest to avoid one adverse outcome. To be conservative, an "adverse outcome" was defined as a patient's failure to achieve at least a 50% improvement in the ODQ at the one-week follow-up. Improvements of smaller magnitudes certainly do not represent an "adverse outcome" in the classic sense and likely still represent clinically meaningful change. Using a more conservative definition of "adverse outcome" as the failure to achieve the MCID of 6 points would yield an even smaller NNT.

The NNT statistic based on the patient's status with respect to the CPR is reported in Table 36. Including all patients who received spinal manipulation (i.e. ignoring the CPR), the NNT was 3.1 (2.2, 5.7) (Table 36). However, when considering only patients classified as positive on the CPR, the NNT drops to 1.3 (1.1, 1.9) (Table 36). On the other hand, considering only patients classified as negative on the CPR, the NNT rises to 9.6 (3.9, Infinity) (Table 36). Similar NNT statistics are observed at the four-week follow-up (Table 36). This means that for every patient a

clinician sees, one adverse outcome is prevented each time the CPR is used. When the CPR is ignored and patients are randomly manipulated, 3-4 patients need to be treated to avoid a patient's not achieving at least a 50% improvement in their ODQ. Incidentally, the confidence intervals around the NNT between all patients who received spinal manipulation (i.e. the heterogeneous group) versus only those classified as positive on the CPR (i.e. the homogeneous group) do not overlap, thus the NNT statistics are statistically different, further elucidating the value of classifying patients with LBP to improve decision-making.

6.9.2.1 Simple to Use

The spinal manipulation CPR is simple to use. Only two out of five criteria in the CPR are based on the results of the physical examination (segmental mobility testing of the lumbar spine and assessment of hip internal rotation range of motion). The other three criteria related to the duration and location of symptoms and score on the FABQW subscale are all obtained during the patient's history. In essence, clinicians can get a good initial impression about whether a patient may benefit from spinal manipulation before the physical examination even begins.

Determining a patient's status with respect to the CPR should take no longer than five minutes, which offers clinicians an efficient and practical evidence-based guide for decision-making to identify patients with LBP likely to benefit from this intervention. Use of the CPR for decision-making in patients with LBP certainly appears to be a valid alternative approach compared to performing a time-consuming plethora of diagnostic tests with little evidence for their use. Importantly, clinicians can also use the prognostic information provided by the LRs associated with the CPR to help patients make informed decisions about potential treatment options for their LBP.

6.9.2.2 One Manipulative Intervention

The manipulative intervention used in the development[12] and validation of the spinal manipulation CPR was chosen based on clinical experience and evidence from the literature[4,5] that it seems to be helpful for a spectrum of patients with LBP. The technique itself is well-described in the literature, easy to perform, and arguably has more evidence for its effectiveness than any other single technique.[4,5,12] The results of this validation study should encourage clinicians that they can be familiar with only one manipulative intervention and still help many patients with LBP. Based on our experience with entry-level physical therapy students, this technique can be easily learned and safely applied by all clinicians. To our knowledge, there are no adverse events that have ever been reported in the literature using this technique in patients with non-radicular LBP.

6.10 The Ultimate Goal: Changing Clinician Behavior to Improve Outcomes of Care

6.11 Level of Evidence of the Spinal Manipulation Clinical Prediction Rule

Although results from this study serve as a necessary step in the CPR's validation, further research is necessary to determine the impact of implementation of the CPR on clinical practice. The validation process of a CPR may ultimately require several studies to fully test its accuracy.[243] However, a single validation study that satisfies four rigorous methodologic standards outlined by McGinn et al[243] may be sufficient to warrant broad implementation. The four standards are listed in Table 42.

Table 42. Methodologic standards for validation of a CPR.

1. Were the patients chosen in an unbiased fashion and do they represent a wide spectrum of

	severity of disease?
2.	Was there a blinded assessment of the criterion standard for all patients?
3.	Was there an explicit and accurate interpretation of the predictor variables and the actual rule without knowledge of the outcome?
4.	Was there 100% follow-up of those enrolled?

This study clearly appears to satisfy each of these criteria. Consecutive patients who met the inclusion/exclusion criteria were enrolled into the study. Except for patients with lower levels of disability or who had neurologic signs or other red flags that might preclude the use of spinal manipulation, all patients with LBP were invited to participate. The reference criterion consisted of a patient self-report measure of outcome using the ODQ, thus not readily subject to rater bias. Patients were unaware of their status with respect to the CPR. Additionally, all predictor variables were assessed at baseline by an examiner who was not aware of which predictor variables were included in the CPR. Furthermore, the patient's overall status with respect to the CPR was made by an examiner blinded to the patient's group assignment. Even presuming the examiner was not blinded to the predictor variables, the reference criterion was not assessed until the one-week follow-up, thus could not foreknow the outcome at baseline when the predictor variables were assessed. Finally, 100% of patients who received spinal manipulation were present for the one-week follow-up, which was the primary follow-up necessary to validate the CPR.

This study was conducted using 13 physical therapists across 8 clinical sites in a variety of healthcare settings and geographical regions in the United States, thus increasing the generalizability of the findings. Table 8 includes a summary of the characteristics of therapists who participated, and Table 9 summarizes the sources from which therapists received their training in spinal manipulation.

McGinn et al[243] have established a hierarchy of evidence for CPRs. Having satisfied rigorous methodologic standards and in light of the ease with which the CPR can be applied and manipulative intervention can be performed, the large effect of treatment, and the extremely low risks associated with spinal manipulation,[124,173,174,176] the spinal manipulation CPR corresponds to Level II in the hierarchy of evidence. Based on these criteria, a recommendation to implement the spinal manipulation CPR into clinical practice seems reasonable.

6.11.1 Impact Analysis of the Spinal Manipulation CPR

The process of developing and testing a CPR requires three steps.[195] Flynn et al[12] accomplished the first step by creating the CPR. The present study addresses the second step in the validation of a CPR. Because the results support the validity of the CPR, the next step will involve an impact analysis of its implementation. This can be done primarily in one of two ways. Ideally, investigators would randomly assign clinical sites to either apply the CPR or not apply it, monitoring the impact of its introduction on clinical practice patterns, outcomes, and costs of care. A design could be utilized in which individual patients were randomly assigned; however, it would be easier for clinicians to incorporate (or not incorporate) the CPR for all patients. An alternative design would be to examine similar outcomes prior to the CPR's implementation and then re-examine the outcomes after it has been implemented. However, the inference of the findings is clearly stronger with the randomized design.

Several successful impact analysis studies[319-321] similar to the one that would be proposed to assess the impact of the spinal manipulation CPR have been completed for the Ottawa ankle rules, making it a Level I CPR in the hierarchy of evidence.[243] One trial[319] randomly assigned 6

emergency departments to apply or not apply the Ottawa ankle rules. Ankle radiographs were ordered in 99.6% of patients in the control group compared to 78.9% in the intervention group (p=.03). Although three fractures were missed, none were associated with an adverse outcome. Utilizing a non-randomized before and after design, Stiell et al[320] demonstrated a 28% reduction in the utilization of ankle radiographs and a 14% reduction in foot radiographs upon implementation of the Ottawa ankle rules compared to a control hospital not trained to use the rule (p<.001). Compared to patients who had radiography but were determined not to have a fracture, patients discharged without radiography also spent significantly less time in the emergency department (80 minutes vs. 116 minutes, p<.0001), had lower estimated total medical costs ($62 vs. $173, p<.001), but did not differ in the percentage that was satisfied with their care (95% vs. 96%). Importantly, these results were achieved without compromising the quality of care, and the reductions were maintained over a 12-month period after the formal trial to assess the impact of the rule was completed.[321] Similar reductions in utilization, costs of care, and waiting times without compromising patient satisfaction or quality of care were found upon implementation of the Ottawa knee rules.[322,323] One could also assess the validity of the CPR in different healthcare settings (i.e. academic medical center vs. military vs. HMO setting) to determine if the rule can be applied across different settings in which healthcare is delivered. The utility of the CPR would be further enhanced if it could be demonstrated that patients benefited similarly when the rule was applied in a broad spectrum of healthcare settings. If these studies were ultimately successful, the spinal manipulation CPR could eventually be classified as a Level I CPR in the hierarchy of evidence.[243]

6.11.2 Implementation Strategies

CPRs have the potential to improve outcomes, increase patient satisfaction, and decrease costs of care. They can be useful tools to save clinicians valuable time and better inform patients as to their diagnosis or prognosis of outcome. It seems logical that publishing evidence for a CPR or a particular intervention should be sufficient to change practice patterns and decision-making accordingly. However, this clearly does not appear to be the case. Although publishing evidence is certainly an important goal, changing clinician behavior is entirely another issue.[324]

Despite the intuitive attraction of CPRs, the transition from evidence to everyday clinical practice can be difficult. The challenge for clinicians is to find an effective means to implement them into a busy clinical setting. Clinicians are required to recall the individual predictor variables, how to assess patients with respect to each predictor, and remember them in the overall decision-making process to maximize the accuracy of its use. Unless clinicians are confident that the CPR is easy to use and will improve costs and/or outcomes of care, systematic implementation may be difficult.

Even having a Level I CPR such as the Ottawa ankle rules does not guarantee that it can be easily incorporated into clinical practice. Cameron and Naylor[325] found no change in the use of ankle radiography among emergency department physicians who had been trained in its use. Although a "magic bullet" strategy to change clinician behavior does not appear to exist,[324] efforts should be made to utilize effective implementation strategies to facilitate practice patterns becoming more consistent with the evidence.[326-328] To achieve successful implementation of the

spinal manipulation CPR into clinical practice, specific strategies will need to be employed based on the unique circumstances of each therapist and practice setting.

Evidence has shown that patients referred early after symptom onset tend to return to worker sooner than patients in which referral is delayed, suggesting that the timing of physical therapy intervention is an important consideration in general.[329] In this study, patients with a shorter duration of symptoms tended to succeed with spinal manipulation; however, only 35.1% (46/131) of patients were below the cut-off of 16 days. The low percentage of patients who met this criterion may largely be attributed to the fact that these patients were all referred for physical therapy from their primary care provider, resulting in a delay of several days or weeks between their referral and initial physical therapy visit. Although not all patients will seek primary management of their LBP within a 16-day period from the time of onset, applying the CPR to patients soon after symptom onset will increase the opportunity for patients to satisfy this criterion. Patients not seen until 2-3 weeks or more later after the onset of their symptoms must satisfy the remaining four criteria in the CPR to meet the 4/5 threshold, resulting in a decreased opportunity to be positive on the CPR. This suggests that efforts need to be made to improve access to physical therapy for patients with acute LBP.

It would also be interesting to determine the impact of implementing the CPR for decision-making among patients with LBP in the primary care setting, where patients first encounter the healthcare system, on practice patterns, outcomes of care, and costs. Although limited evidence exists, one study[330] demonstrated that training primary care physicians in a limited number of manual therapy interventions may result in improved recovery rates immediately after treatment.

Given the parochial boundaries that exist among professions who provide similar services, such an endeavor would likely evoke controversy. However, spinal manipulation is not the exclusive domain of any single profession. Physical therapists, medical doctors, osteopathic physicians, and chiropractors all include these skills in their scope of practice, and evidence-based practice should have no professional boundaries. Rather, maximizing patient outcome and quality of life must remain the primary focus, without undue regard for advancing any single profession's political agenda.

6.11.3 General vs. Specific Approach

A recent RCT[193] was performed to assess the meaningful of end-feel testing to improve decision-making in the use of manipulation for patients with neck pain. 104 patients were randomly assigned to receive one of two interventions. One group received manipulation targeted to specific cervical vertebrae based on the results of precise end-feel testing. For the other group, end-feel testing was performed to rule out the possibility of an attention effect, but decision-making was based on randomly computer-generated examination findings. Although both groups improved, no differences in neck pain or stiffness were found between the groups five hours after treatment. Although short-term, these results suggest that improvement from manipulation may be attributable to the manipulative intervention itself, rather than the explicit decision-making process that is used. In a similar study in patients with LBP, it would be interesting to determine if using a generalized manipulative intervention in patients who meet the criteria in the CPR results in similar or perhaps even better outcome than using some of the highly specific and often complex diagnostic schemes theoretically used to select a specific manipulative intervention to ameliorate a specific underlying biomechanical dysfunction. Additionally, a study could be conducted to determine if there are differences in the effectiveness of a variety of manipulative

interventions in patients who meet the criteria in the CPR. Perhaps the use of manipulation itself in the appropriate subgroup of patients with LBP is more important than the selection of a precise technique based on theoretical principles about biomechanical dysfunction that have largely not been substantiated by the evidence.

7. Conclusion

CPRs have become increasingly important to improve decision-making related to diagnosis and prognosis among patients with a variety of disorders. To our knowledge, the development and validation of the spinal manipulation CPR is the first to establish a patient's prognosis based on receiving a standardized intervention. Additionally, its development and validation are the first to examine the characteristics of patients most likely to benefit from this intervention. The results of this study validate and refine the initial development of the spinal manipulation CPR and suggest that outcome from spinal manipulation depends upon a patient's status with respect to the CPR. Importantly, the criteria in the CPR appear to identify patients specifically responding to spinal manipulation, rather than simply identifying patients with a favorable natural history.

The development and validation of the spinal manipulation CPR, combined with the increased risk of worsening when this intervention is not offered to patients, should help to reverse trends showing spinal manipulation continues to be underutilized despite consistent recommendations for its use. The results of this study should also encourage the inclusion of these skills in entry-level curricula, ultimately resulting in increased utilization of these skills among entry-level therapists. Patients who might benefit from this intervention can be accurately identified at the initial examination by assessing only a few factors from history and physical examination. Because the CPR is simple to use and requires proficiency with only a single manipulative

intervention, it can readily be applied in a busy clinical setting by clinicians with varying levels of experience.

Future studies will continue to validate the spinal manipulation CPR and examine the impact of its implementation on clinical practice patterns, outcomes, and costs of care. Future work from this study will also investigate outcomes from spinal manipulation at a 6-month follow-up. Armed with only a single manipulative intervention, clinicians can use the spinal manipulation CPR to improve decision-making and outcome for patients with LBP.

8. Appendices

8.1 APPENDIX A

Lumbopelvic Region Diagnostic Test Operational Definitions

8.2 APPENDIX B

Lumbopelvic Region Diagnostic Tests: Relationship to Success with Spinal Manipulation

8.3 APPENDIX C

Screening Examination

8.4 APPENDIX D

Patient Eligibility Tracking

8.5 APPENDIX E

Manual of Standard Operations and Procedures (MSOP)

8.6 APPENDIX F

Demographic Information

8.7 APPENDIX G

Pain Diagram and Rating

8.8 APPENDIX H

Fear-Avoidance Beliefs Questionnaire (FABQ)

8.9 APPENDIX I

Oswestry Disability Questionnaire (ODQ)

8.10 APPENDIX J

Physical Examination Form

8.11 APPENDIX K

Manipulation Group Exercise Program (Sessions #1-2 only)

8.12 APPENDIX L

Treatment Form – Manipulation Group

8.13 APPENDIX M

Theoretical Rational for Exercise Program

8.14 APPENDIX N

Exercise Group (Sessions #1-5) and Manipulation Group (Sessions #3-5)

8.15 APPENDIX O

Treatment Form – Exercise Group

9. BIBLIOGRAPHY

1. Postacchini F, Facchini M, Palieri P. Efficacy of various forms of conservative treatments in low back pain: a comparative study. *Neuro Orthop*. 1988; 6:28-35.

2. Rasmussen GG. Manipulation in treatment of low-back pain: a randomized clinical trial. *Manual Med.* 1979; 1:8-10.

3. Wreje U, Nordgren B, Aberg H. Treatment of pelvic joint dysfunction in primary care--a controlled study. *Scand J Prim Health Care*. 1992; 10(4):310-315.

4. Delitto A, Cibulka MT, Erhard RE, Bowling RW, Tenhula JA. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. *Phys Ther*. 1993; 73(4):216-222.

5. Erhard RE, Delitto A, Cibulka MT. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. *Phys Ther*. 1994; 74(12):1093-1100.

6. Koes BW, Bouter LM, van Mameren H, Essers AH, Verstegen GM, Hofhuizen DM et al. The effectiveness of manual therapy, physiotherapy, and treatment by the general practitioner for nonspecific back and neck complaints. A randomized clinical trial. *Spine*. 1992; 17(1):28-35.

7. Triano JJ, McGregor M, Hondras MA, Brennan PC. Manipulative therapy versus education programs in chronic low back pain. *Spine*. 1995; 20(8):948-955.

8. Cherkin DC, Deyo RA, Battie M, Street J, Barlow W. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. *N Engl J Med.* 1998; 339(15):1021-1029.

9. Gibson T, Grahame R, Harkness J, Woo P, Blagrave P, Hills R. Controlled comparison of short-wave diathermy treatment with osteopathic treatment in non-specific low back pain. *Lancet.* 1985; 1(8440):1258-1261.

10. Godfrey CM, Morgan PP, Schatzker J. A randomized trial of manipulation for low-back pain in a medical setting. *Spine.* 1984; 9(3):301-304.

11. Glover JR, Morris JG, Khosla T. Back pain: a randomized clinical trial of rotational manipulation of the trunk. *Br J Ind Med.* 1974; 31(1):59-64.

12. Flynn T, Fritz J, Whitman J, Wainner R, Magel J, Butler B et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short term improvement with spinal manipulation. *Spine.* 2002; 27(24):2835-2843.

13. Deyo RA, Phillips WR. Low back pain. A primary care challenge. *Spine.* 1996; 21(24):2826-2832.

14. Patel AT, Ogle AA. Diagnosis and management of acute low back pain. *Am Fam Physician.* 2000; 61(6):1779-1790.

15. Hart LG, Deyo RA, Cherkin DC. Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine.* 1995; 20(1):11-19.

16. Webster BS, Snook SH. The cost of compensable low back pain. *J Occup Med.* 1990; 32(1):13-15.

17. van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. *Pain.* 1995; 62(2):233-240.

18. van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine.* 1997; 22(18):2128-2156.

19. van Tulder MW, Malmivaara A, Esmail R, Koes BW. Exercise therapy for low back pain. *Cochrane Database Syst Rev.* 2000;(2):CD000335.

20. Philadelphia Panel Evidence-Based Clinical Practice Guidelines on Selected Rehabilitation Interventions for Low Back Pain. *Phys Ther.* 2001; 81(10):1641-1674.

21. Bigos S, Bowyer O, Braen G, Brown K, Deyo R, Haldeman S et al. *Acute Low Back Problems in Adults.* AHCPR Publication 95-0642. 1994. Rockville,MD, Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services.

22. Faas A. Exercises: which ones are worth trying, for which patients, and when? *Spine.* 1996; 21(24):2874-2878.

23. Bouter LM, van Tulder MW, Koes BW. Methodologic issues in low back pain research in primary care. *Spine.* 1998; 23(18):2014-2020.

24. Leboeuf-Yde C, Lauritsen JM, Lauritzen T. Why has the search for causes of low back pain largely been nonconclusive? *Spine.* 1997; 22(8):877-881.

25. Dreyfuss P, Michaelsen M, Pauza K, McLarty J, Bogduk N. The value of medical history and physical examination in diagnosing sacroiliac joint pain. *Spine.* 1996; 21(22):2594-2602.

26. Schwarzer AC, Aprill CN, Bogduk N. The sacroiliac joint in chronic low back pain. *Spine.* 1995; 20(1):31-37.

27. Sturesson B, Selvik G, Uden A. Movements of the sacroiliac joints. A roentgen stereophotogrammetric analysis. *Spine.* 1989; 14(2):162-165.

28. Sturesson B, Uden A, Vleeming A. A radiostereometric analysis of movements of the sacroiliac joints during the standing hip flexion test. *Spine.* 2000; 25(3):364-368.

29. Sturesson B, Uden A, Vleeming A. A radiostereometric analysis of the movements of the sacroiliac joints in the reciprocal straddle position. *Spine.* 2000; 25(2):214-217.

30. Tullberg T, Blomberg S, Branth B, Johnsson R. Manipulation does not alter the position of the sacroiliac joint. A roentgen stereophotogrammetric analysis. *Spine.* 1998; 23(10):1124-1128.

31. *NIOSH Low Back Atlas of Standardized Tests and Measures.* Morgantown, WV: National Institute for Occupational Safety and Health, Division of Safety Research, 1998.

32. van der WP, Hagmeijer RH, Meyne W. Clinical tests of the sacroiliac joint. A systemic methodological review. Part 1: reliability. *Man Ther.* 2000; 5(1):30-36.

33. Hestbaek L, Leboeuf-Yde C. Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review. *J Manipulative Physiol Ther.* 2000; 23(4):258-275.

34. Freburger JK, Riddle DL. Using published evidence to guide the examination of the sacroiliac joint region. *Phys Ther.* 2001; 81(5):1135-1143.

35. van der WP, Meyne W, Hagmeijer RH. Clinical tests of the sacroiliac joint. *Man Ther.* 2000; 5(2):89-96.

36. Ebrall P. Motion palpation: It's time to accept the evidence. *J Manipulative Physiol Ther.* 2000; 23(7):513-514.

37. Harrison DE, Harrison DD, Troyanovich SJ, Harmon S. A normal spinal position: It's time to accept the evidence. *J Manipulative Physiol Ther.* 2000; 23(9):623-644.

38. Alviso DJ, Dong GT, Lentell GL. Intertester reliability for measuring pelvic tilt in standing. *Phys Ther.* 1988; 68(9):1347-1351.

39. Crowell RD, Cummings GS, Walker JR, Tillman LJ. Intratester and intertester reliability and validity of measures of innominate bone inclination. *J Orthop Sports Phys Ther.* 1994; 20(2):88-97.

40. Laslett M, Williams M. The reliability of selected pain provocation tests for sacroiliac joint pathology. *Spine.* 1994; 19(11):1243-1249.

41. McKinnon R, Balkcum M. *Inter-rater reliability of combined observation and palpation of sitting iliac crest height.* 8, 135-136. 2000.

42. Levangie PK. Four clinical tests of sacroiliac joint dysfunction: the association of test results with innominate torsion among patients with and without low back pain. *Phys Ther.* 1999; 79(11):1043-1057.

43. Levangie PK. The association between static pelvic asymmetry and low back pain. *Spine.* 1999; 24(12):1234-1242.

44. Maigne JY, Aivaliklis A, Pfefer F. Results of sacroiliac joint double block and value of sacroiliac pain provocation tests in 54 patients with low back pain. *Spine.* 1996; 21(16):1889-1892.

45. Potter NA, Rothstein JM. Intertester reliability for selected clinical tests of the sacroiliac joint. *Phys Ther.* 1985; 65(11):1671-1675.

46. Herzog W, Read LJ, Conway PJ, Shaw LD, McEwen MC. Reliability of motion palpation procedures to detect sacroiliac joint fixations. *J Manipulative Physiol Ther.* 1989; 12(2):86-92.

47. Strender LE, Sjoblom A, Sundell K, Ludwig R, Taube A. Interexaminer reliability in physical examination of patients with low back pain. *Spine.* 1997; 22(7):814-820.

48. Hawk C, Phongphua C, Bleecker J, Swank L, Lopez D, Rubley T. Preliminary study of the reliability of assessment procedures for indications for chiropractic adjustments of the lumbar spine. *J Manipulative Physiol Ther.* 1999; 22(6):382-389.

49. Meijne W, van Neerbos K, Aufdemkampe G, van der WP. Intraexaminer and interexaminer reliability of the Gillet test. *J Manipulative Physiol Ther.* 1999; 22(1):4-9.

50. Vincent-Smith B, Gibbons P. Inter-examiner and intra-examiner reliability of the standing flexion test. *Man Ther.* 1999; 4(2):87-93.

51. O'Haire C, Gibbons P. Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study. *Man Ther.* 2000; 5(1):13-20.

52. Riddle DL, Freburger JK. Evaluation of the presence of sacroiliac joint region dysfunction using a combination of tests: a multicenter intertester reliability study. *Phys Ther.* 2002; 82(8):772-781.

53. French SD, Green S, Forbes A. Reliability of chiropractic methods commonly used to detect manipulable lesions in patients with chronic low-back pain. *J Manipulative Physiol Ther.* 2000; 23(4):231-238.

54. Freburger JK, Riddle DL. Measurement of sacroiliac joint dysfunction: a multicenter intertester reliability study. *Phys Ther.* 1999; 79(12):1134-1141.

55. Keating JC. Several strategies for evaluating the objectivity of measurements in clinical research abd practice. *Journal of the Canadian Chiropractic Association.* 1988; 32:133-138.

56. Gerlach UJ, Lierse W. Functional construction of the sacroiliac ligamentous apparatus. *Acta Anat (Basel).* 1992; 144(2):97-102.

57. Cramer GD, Gregerson DM, Knudsen JT, Hubbard BB, Ustas LM, Cantu JA. The effects of side-posture positioning and spinal adjusting on the lumbar Z joints: a randomized controlled trial with sixty-four subjects. *Spine.* 2002; 27(22):2459-2466.

58. Childs JD, Piva SR, Erhard RE, Hicks GE. Side-to-side weight-bearing asymmetry in patients with low back pain. *Manual Therapy.* 2003; 8(3):166-169.

59. Childs JD, Piva SR, Erhard RE. Immediate improvements in side-to-side weight-bearing asymmetry and iliac crest symmetry after manipulation in patients with low back pain. *J Manipulative Physiol Ther.* In press.

60. Broadhurst NA, Bond MJ. Pain provocation tests for the assessment of sacroiliac joint dysfunction. *J Spinal Disord.* 1998; 11(4):341-345.

61. Slipman CW, Sterenfeld EB, Chou LH, Herzog R, Vresilovic E. The predictive value of provocative sacroiliac joint stress maneuvers in the diagnosis of sacroiliac joint syndrome. *Arch Phys Med Rehabil.* 1998; 79(3):288-292.

62. Cibulka MT, Koldehoff R. Clinical usefulness of a cluster of sacroiliac joint tests in patients with and without low back pain. *J Orthop Sports Phys Ther.* 1999; 29(2):83-89.

63. Schwarzer AC, Aprill CN, Derby R, Fortin J, Kine G, Bogduk N. The false-positive rate of uncontrolled diagnostic blocks of the lumbar zygapophysial joints. *Pain.* 1994; 58(2):195-200.

64. Hogan QH, Abram SE. Neural blockade for diagnosis and prognosis. A review. *Anesthesiology.* 1997; 86(1):216-241.

65. Barnsley L, Lord S, Wallis B, Bogduk N. False-positive rates of cervical zygapophysial joint blocks. *Clin J Pain.* 1993; 9(2):124-130.

66. Childs JD, Fritz JM, Piva SR, Erhard RE. Clinical decision-making in the identification of patients likely to benefit from spinal manipulation: A traditional versus and evidence-based approach. *J Orthop Sports Phys Ther.* 2003; 33(5):259-272.

67. Erhard RE, Bowling RW. The recognition and treatment of the pelvic component of low back and sciatic pain. *Bulletin of the Orthopaedic Section, American Physical Therapy Association.* 1977; 2(3):4-15.

68. American Physical Therapy Association (APTA). *Guide to Physical Therapist Practice.* 2nd ed. APTA, 2001.

69. Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM. Spinal manipulation for low back pain. An updated systematic review of randomized clinical trials. *Spine.* 1996; 21(24):2860-2871.

70. Ottenbacher K, DiFabio RP. Efficacy of spinal manipulation/mobilization therapy. A meta-analysis. *Spine.* 1985; 10(9):833-837.

71. DonTigny R. Dysfunction of the sacroiliac joint and its treatment. *J Orthop Sports Phys Ther.* 1979; 1:23-29.

72. Hawk C, Azad A, Phongphua C, Long CR. Preliminary study of the effects of a placebo chiropractic treatment with sham adjustments. *J Manipulative Physiol Ther.* 1999; 22(7):436-443.

73. Meade TW, Dyer S, Browne W, Townsend J, Frank AO. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *BMJ.* 1990; 300(6737):1431-1437.

74. Sims-Williams H, Jayson MI, Young SM, Baddeley H, Collins E. Controlled trial of mobilisation and manipulation for patients with low back pain in general practice. *Br Med J.* 1978; 2(6148):1338-1340.

75. Evans DP, Burke MS, Lloyd KN, Roberts EE, Roberts GM. Lumbar spinal manipulation on trial. Part I--clinical assessment. *Rheumatol Rehabil.* 1978; 17(1):46-53.

76. Sims-Williams H, Jayson MI, Young SM, Baddeley H, Collins E. Controlled trial of mobilisation and manipulation for low back pain: hospital patients. *Br Med J.* 1979; 2(6201):1318-1320.

77. Hoehler FK, Tobis JS, Buerger AA. Spinal manipulation for low back pain. *JAMA.* 1981; 245(18):1835-1838.

78. Farrell JP, Twomey LT. Acute low back pain. Comparison of two conservative treatment approaches. *Med J Aust.* 1982; 1(4):160-164.

79. Nwuga VC. Relative therapeutic efficacy of vertebral manipulation and conventional treatment in back pain management. *Am J Phys Med.* 1982; 61(6):273-278.

80. Hadler NM, Curtis P, Gillings DB, Stinnett S. A benefit of spinal manipulation as adjunctive therapy for acute low-back pain: a stratified controlled trial. *Spine.* 1987; 12(7):702-706.

81. Mathews JA, Mills SB, Jenkins VM, Grimes SM, Morkel MJ, Mathews W et al. Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *Br J Rheumatol.* 1987; 26(6):416-423.

82. Ongley MJ, Klein RG, Dorman TA, Eek BC, Hubert LJ. A new approach to the treatment of chronic low back pain. *Lancet.* 1987; 2(8551):143-146.

83. Mathews JA. Manipulation and traction for sciatica and lumbago: Physiotherapeutic techniques in two controlled trials. *Physiotherapy Practice.* 1987; 4:201-206.

84. MacDonald RS, Bell CM. An open controlled assessment of osteopathic manipulation in nonspecific low-back pain. *Spine.* 1990; 15(5):364-370.

85. Sanders GE, Reinert O, Tepe R, Maloney P. Chiropractic adjustive manipulation on subjects with acute low back pain: visual analog pain scores and plasma beta-endorphin levels. *J Manipulative Physiol Ther.* 1990; 13(7):391-395.

86. Blomberg S, Svardsudd K, Mildenberger F. A controlled, multicentre trial of manual therapy in low-back pain. Initial status, sick-leave and pain score during follow-up. *Scand J Prim Health Care.* 1992; 10(3):170-178.

87. Hsieh CY, Phillips RB, Adams AH, Pope MH. Functional outcomes of low back pain: comparison of four treatment groups in a randomized controlled trial. *J Manipulative Physiol Ther.* 1992; 15(1):4-9.

88. Koes BW, Bouter LM, van Mameren H, Essers AH, Verstegen GM, Hofhuizen DM et al. Randomised clinical trial of manipulative therapy and physiotherapy for persistent back and neck complaints: results of one year follow up. *BMJ.* 1992; 304(6827):601-605.

89. Blomberg S, Svardsudd K, Tibblin G. Manual therapy with steroid injections in low-back pain. Improvement of quality of life in a controlled trial with four months' follow-up. *Scand J Prim Health Care.* 1993; 11(2):83-90.

90. Koes BW, Bouter LM, van Mameren H, Essers AH, Verstegen GJ, Hofhuizen DM et al. A randomized clinical trial of manual therapy and physiotherapy for persistent back and neck complaints: subgroup analysis and relationship between outcome measures. *J Manipulative Physiol Ther.* 1993; 16(4):211-219.

91. Blomberg S, Hallin G, Grann K, Berg E, Sennerby U. Manual therapy with steroid injections--a new approach to treatment of low back pain. A controlled multicenter trial with an evaluation by orthopedic surgeons. *Spine*. 1994; 19(5):569-577.

92. Blomberg S, Svardsudd K, Tibblin G. A randomized study of manual therapy with steroid injections in low-back pain. Telephone interview follow-up of pain, disability, recovery and drug consumption. *Eur Spine J*. 1994; 3(5):246-254.

93. Hurwitz EL. The relative impact of chiropractic vs. medical management of low back pain on health status in a multispecialty group practice. *J Manipulative Physiol Ther*. 1994; 17(2):74-82.

94. Meade TW, Dyer S, Browne W, Frank AO. Randomised comparison of chiropractic and hospital outpatient management for low back pain: results from extended follow up. *BMJ*. 1995; 311(7001):349-351.

95. Bronfort G, Goldsmith CH, Nelson CF, Boline PD, Anderson AV. Trunk exercise combined with spinal manipulative or NSAID therapy for chronic low back pain: a randomized, observer-blinded clinical trial. *J Manipulative Physiol Ther*. 1996; 19(9):570-582.

96. Seferlis T, Nemeth G, Carlsson AM, Gillstrom P. Conservative treatment in patients sick-listed for acute low-back pain: a prospective randomised study with 12 months' follow-up. *Eur Spine J*. 1998; 7(6):461-470.

97. Giles LG, Muller R. Chronic spinal pain syndromes: a clinical pilot trial comparing acupuncture, a nonsteroidal anti-inflammatory drug, and spinal manipulation. *J Manipulative Physiol Ther*. 1999; 22(6):376-381.

98. Burton AK, Tillotson KM, Cleary J. Single-blind randomised controlled trial of chemonucleolysis and manipulation in the treatment of symptomatic lumbar disc herniation. *Eur Spine J.* 2000; 9(3):202-207.

99. Nyiendo J, Haas M, Goodwin P. Patient characteristics, practice activities, and one-month outcomes for chronic, recurrent low-back pain treated by chiropractors and family medicine physicians: a practice-based feasibility study. *J Manipulative Physiol Ther.* 2000; 23(4):239-245.

100. Hemmila HM, Keinanen-Kiukaanniemi SM, Levoska S, Puska P. Long-term effectiveness of bone-setting, light exercise therapy, and physiotherapy for prolonged back pain: a randomized controlled trial. *J Manipulative Physiol Ther.* 2002; 25(2):99-104.

101. Hertzman-Miller RP, Morgenstern H, Hurwitz EL, Yu F, Adams AH, Harber P et al. Comparing the satisfaction of low back pain patients randomized to receive medical or chiropractic care: results from the UCLA low-back pain study. *Am J Public Health.* 2002; 92(10):1628-1633.

102. Aure OF, Hoel NJ, Vasseljen O. Manual Therapy and Exercise Therapy in Patients With Chronic Low Back Pain: A Randomized, Controlled Trial With 1-Year Follow-up. *Spine.* 2003; 28(6):525-531.

103. Doran DM, Newell DJ. Manipulation in treatment of low back pain: a multicentre study. *Br Med J.* 1975; 2(5964):161-164.

104. Jayson MI, Sims-Williams H, Young S, Baddeley H, Collins E. Mobilization and manipulation for low-back pain. *Spine.* 1981; 6(4):409-416.

105. Zylbergold RS, Piper MC. Lumbar disc disease: comparative analysis of physical therapy treatments. *Arch Phys Med Rehabil.* 1981; 62(4):176-179.

106. Coxhead CE, Inskip H, Meade TW, North WR, Troup JD. Multicentre trial of physiotherapy in the management of sciatic symptoms. *Lancet.* 1981; 1(8229):1065-1068.

107. Waterworth RF, Hunter IA. An open study of diflunisal, conservative and manipulative therapy in the management of acute mechanical low back pain. *N Z Med J.* 1985; 98(779):372-375.

108. Herzog W, Conway PJ, Willcox BJ. Effects of different treatment modalities on gait symmetry and clinical measures for sacroiliac joint patients. *J Manipulative Physiol Ther.* 1991; 14(2):104-109.

109. Pope MH, Phillips RB, Haugh LD, Hsieh CY, MacDonald L, Haldeman S. A prospective randomized three-week trial of spinal manipulation, transcutaneous muscle stimulation, massage and corset in the treatment of subacute low back pain. *Spine.* 1994; 19(22):2571-2577.

110. Timm KE. A randomized-control study of active and passive treatments for chronic low back pain following L5 laminectomy. *J Orthop Sports Phys Ther.* 1994; 20(6):276-286.

111. Skargren EI, Oberg BE, Carlsson PG, Gade M. Cost and effectiveness analysis of chiropractic and physiotherapy treatment for low back and neck pain. Six-month follow-up. *Spine.* 1997; 22(18):2167-2177.

112. Hemmila HM, Keinanen-Kiukaanniemi SM, Levoska S, Puska P. Does folk medicine work? A randomized clinical trial on patients with prolonged back pain. *Arch Phys Med Rehabil.* 1997; 78(6):571-577.

113. Skargren EI, Carlsson PG, Oberg BE. One-year follow-up comparison of the cost and effectiveness of chiropractic and physiotherapy as primary management for back pain. Subgroup analysis, recurrence, and additional health care utilization. *Spine.* 1998; 23(17):1875-1883.

114. Andersson GB, Lucente T, Davis AM, Kappler RE, Lipton JA, Leurgans S. A comparison of osteopathic spinal manipulation with standard care for patients with low back pain. *N Engl J Med.* 1999; 341(19):1426-1431.

115. Hsieh CY, Adams AH, Tobis J, Hong CZ, Danielson C, Platt K et al. Effectiveness of four conservative treatments for subacute low back pain: a randomized clinical trial. *Spine.* 2002; 27(11):1142-1148.

116. Hurwitz EL, Morgenstern H, Harber P, Kominski GF, Belin TR, Yu F et al. A randomized trial of medical care with and without physical therapy and chiropractic care with and without physical modalities for patients with low back pain: 6-month follow-up outcomes from the UCLA low back pain study. *Spine.* 2002; 27(20):2193-2204.

117. Moritz U. Evaluation of manipulation and other manual therapy. Criteria for measuring the effect of treatment. *Scand J Rehabil Med.* 1979; 11(4):173-179.

118. Greenland S, Reisbord LS, Haldeman S, Buerger AA. Controlled clinical trials of manipulation: a review and a proposal. *J Occup Med.* 1980; 22(10):670-676.

119. Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ.* 1991; 303(6813):1298-1303.

120. Abenhaim L, Bergeron AM. Twenty years of randomized clinical trials of manipulative therapy for back pain: a review. *Clin Invest Med.* 1992; 15(6):527-535.

121. Anderson R, Meeker WC, Wirick BE, Mootz RD, Kirk DH, Adams A. A meta-analysis of clinical trials of spinal manipulation. *J Manipulative Physiol Ther.* 1992; 15(3):181-194.

122. Assendelft WJ, Koes BW, van der Heijden GJ, Bouter LM. The efficacy of chiropractic manipulation for back pain: blinded review of relevant randomized clinical trials. *J Manipulative Physiol Ther.* 1992; 15(8):487-494.

123. Di Fabio RP. Efficacy of manual therapy. *Phys Ther.* 1992; 72(12):853-864.

124. Shekelle PG, Adams AH, Chassin MR, Hurwitz EL, Brook RH. Spinal manipulation for low-back pain. *Ann Intern Med.* 1992; 117(7):590-598.

125. Bronfort G. Spinal manipulation: current state of research and its indications. *Neurol Clin.* 1999; 17(1):91-111.

126. Mior S. Manipulation and mobilization in the treatment of chronic pain. *Clin J Pain.* 2001; 17(4 Suppl):S70-S76.

127. Di Fabio RP. Clinical assessment of manipulation and mobilization of the lumbar spine. A critical review of the literature. *Phys Ther.* 1986; 66(1):51-54.

128. Beckerman H, Bouter LM, van der Heijden GJ, de Bie RA, Koes BW. Efficacy of physiotherapy for musculoskeletal disorders: what can we learn from research? *Br J Gen Pract.* 1993; 43(367):73-77.

129. Assendelft WJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. *JAMA.* 1995; 274(24):1942-1948.

130. Koes BW, Bouter LM, van der Heijden GJ. Methodological quality of randomized clinical trials on treatment efficacy in low back pain. *Spine.* 1995; 20(2):228-235.

131. Scheer SJ, Radack KL, O'Brien DR, Jr. Randomized controlled trials in industrial low back pain relating to return to work. Part 1. Acute interventions. *Arch Phys Med Rehabil.* 1995; 76(10):966-973.

132. Assendelft WJ, Koes BW, van der Heijden GJ, Bouter LM. The effectiveness of chiropractic for treatment of low back pain: an update and attempt at statistical pooling. *J Manipulative Physiol Ther.* 1996; 19(8):499-507.

133. Brox JI, Hagen KB, Juel NG, Storheim K. Is exercise therapy and manipulation effective in low back pain? *Tidsskr Nor Laegeforen.* 1999; 119(14):2042-2050.

134. Vernon H. Qualitative review of studies of manipulation-induced hypoalgesia. *J Manipulative Physiol Ther.* 2000; 23(2):134-138.

135. Ernst E. Does spinal manipulation have specific treatment effects? *Fam Pract.* 2000; 17(6):554-556.

136. Ernst E, Harkness E. Spinal manipulation. a systematic review of sham-controlled, double-blind, randomized clinical trials. *J Pain Symptom Manage.* 2001; 22(4):879-889.

137. Ferreira ML, Ferreira PH, Latimer J, Herbert R, Maher CG. Does spinal manipulative therapy help people with chronic low back pain? *Aust J Physiother.* 2002; 48(4):277-284.

138. Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG. Spinal manipulative therapy for low back pain. A meta-analysis of effectiveness relative to other therapies. *Ann Intern Med.* 2003; 138(11):871-881.

139. Cherkin DC, Sherman KJ, Deyo RA, Shekelle PG. A review of the evidence for the effectiveness, safety, and cost of acupuncture, massage therapy, and spinal manipulation for back pain. *Ann Intern Med.* 2003; 138(11):898-906.

140. Assendelft WJ, Lankhorst GJ. Effectiveness of manipulative therapy in low back pain: systematic literature reviews and guidelines are inconclusive. *Ned Tijdschr Geneeskd.* 1998; 142(13):684-687.

141. Koes BW, van Tulder MW, Ostelo R, Kim BA, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. *Spine.* 2001; 26(22):2504-2513.

142. Panel E. *Veteran's Health Adminsistration/Department of Defense Clinical Practice Guideline for the Management of Low Back Pain or Sciatica in the Primary Care Setting.* 5-1-1999. Washington, D.C., The Office of Performance and Quality, Veteran's Health Administration and the Quality Management Directorate, United States Army Medical Command.

143. ACC tNHC. *New Zealand Acute Low Back Pain Guide.* 1997. Wellington, New Zealand.

144. Manniche C. *Low back pain: Frequency, management, and prevention from HTA perspective.* 1. 1999. Danish Institute for Health Technology Assessment.

145. Malmivaara A, Kotilainen E, Laasonen E, et al. *Clinical Practice Guidelines of the Finnish Medical Association Duodecim: Diseases of the Low Back.* 1999.

146. Waddell G, McIntosh A, Hutchinson A, Feder G, Lewis M. *Clinical Guidelines for the Management of Acute Low Back Pain.* 1999. London, Royal College of General Practitioners.

147. The Swedish Council on Technology Assessment in Health Care. *Neck and Back Pain: The Scientific Evidence of Causes, Diagnosis, and Treatment.* Philadelphia: Lippincott and Wilkins , 2000.

148. Keel P, Weber M, Roux E, et al. *Kreuzschmerzen: Hintergründe, prävention, behandlung.* 1998. Bern, Switzerland, Swiss Medical Society (FMH).

149. Drug Committee of the German Medical Society. Treatment Guideline - Backache. *Z Artzl Fortbild Qualitatssich.* 1997; 91:457-460.

150. Faas A, Chavannes AW, Koes BW, et al. NHG-Standaard Lage-Rugpijn. *Huisarts Wet.* 1996; 39:18-31.

151. Bogduk N. *Draft evidence based clinical guidelines for the management of acute low back pain.* 2000. Australia, National Health and Medical Research Council.

152. Borkan J, Reis S, Werner S, et al. *Guidelines for treating low back pain in primary care.* 1996. Harfuah, Israel, The Israeli Low Back Pain Guideline Group.

153. Moffroid MT, Haugh LD, Henry SM, Short B. Distinguishable groups of musculoskeletal low back pain patients and asymptomatic control subjects based on physical measures of the NIOSH Low Back Atlas. *Spine.* 1994; 19(12):1350-1358.

154. Sikorski JM. A rationalized approach to physiotherapy for low-back pain. *Spine.* 1985; 10(6):571-579.

155. Van Dillen LR, Sahrmann SA, Norton BJ, Caldwell CA, Fleming DA, McDonnell MK et al. Reliability of physical examination items used for classification of patients with low back pain. *Phys Ther.* 1998; 78(9):979-988.

156. McKenzie RA. *The Lumbar Spine: Mechanical Diagnosis and Therapy.* Waikanae, New Zealand: Spinal Publications Limited, 1989.

157. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Phys Ther.* 1995; 75(6):470-485.

158. *Manual of the International Classification of Diseases, Injuries, and Causes of Death.* Geneva, Switzerland: World Health Organization, 1977.

159. American Physical Therapy Association (APTA). Clinical Research Agenda for Physical Therapy. *Physical Therapy.* 2000;(80):499-513.

160. Borkan JM, Koes B, Reis S, Cherkin DC. A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. *Spine.* 1998; 23(18):1992-1996.

161. van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain: a systematic review within the framework of the cochrane collaboration back review group. *Spine.* 2000; 25(21):2784-2796.

162. Rose SJ. Physical therapy diagnosis: role and function. *Phys Ther.* 1989; 69(7):535-537.

163. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine.* 2nd ed. Boston: Little, Brown and Company, 1991.

164. Cherkin DC, Deyo RA, Wheeler K, Ciol MA. Physician views about treating low back pain. The results of a national survey. *Spine.* 1995; 20(1):1-9.

165. Jette DU, Jette AM. Professional uncertainty and treatment choices by physical therapists. *Arch Phys Med Rehabil.* 1997; 78(12):1346-1351.

166. Li LC, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. *Phys Ther.* 2001; 81(4):1018-1028.

167. Gracey JH, McDonough SM, Baxter GD. Physiotherapy management of low back pain: a survey of current practice in northern ireland. *Spine.* 2002; 27(4):406-411.

168. Armstrong MP, McDonough S, Baxter GD. Clinical guidelines versus clinical practice in the management of low back pain. *Int J Clin Pract.* 2003; 57(1):9-13.

169. Delitto A. Clinicians and researchers who treat and study patients with low back pain: are you listening? *Phys Ther.* 1998; 78(7):705-707.

170. Adams G, Sim J. A survey of UK manual therapists' practice of and attitudes towards manipulation and its complications. *Physiother Res Int.* 1998; 3(3):206-227.

171. Hurley L, Yardley K, Gross AR, Hendry L, McLaughlin L. A survey to examine attitudes and patterns of practice of physiotherapists who perform cervical spine manipulation. *Man Ther.* 2002; 7(1):10-18.

172. Hudson-Cook N, Tomes-Nicholson K, Breen A. A Revised Oswestry Disability Questionnaire. In: Roland MO, Jenner JR, editors. *Back Pain: New Approaches to Rehabilitation and Education.* New York: Manchester University Press, 1989: 187-204.

173. Haldeman S, Rubinstein SM. Cauda equina syndrome in patients undergoing manipulation of the lumbar spine. *Spine.* 1992; 17(12):1469-1473.

174. Assendelft WJ, Bouter LM, Knipschild PG. Complications of spinal manipulation: a comprehensive review of the literature. *J Fam Pract.* 1996; 42(5):475-480.

175. Senstad O, Leboeuf-Yde C, Borchgrevink C. Frequency and characteristics of side effects of spinal manipulative therapy. *Spine.* 1997; 22(4):435-440.

176. Leboeuf-Yde C, Hennius B, Rudberg E, Leufvenmark P, Thunman M. Side effects of chiropractic treatment: a prospective study. *J Manipulative Physiol Ther.* 1997; 20(8):511-515.

177. Hungin AP, Kean WF. Nonsteroidal anti-inflammatory drugs: overused or underused in osteoarthritis? *Am J Med.* 2001; 110(1 Suppl 1):S8-S11.

178. Tannenbaum H, Davis P, Russell AS, Atkinson MH, Maksymowych W, Huang SH et al. An evidence-based approach to prescribing NSAIDs in musculoskeletal disease: a Canadian consensus. Canadian NSAID Consensus Participants. *CMAJ.* 1996; 155(1):77-88.

179. Gabriel SE, Jaakkimainen L, Bombardier C. Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs. A meta-analysis. *Ann Intern Med.* 1991; 115(10):787-796.

180. Tamblyn R, Berkson L, Dauphinee WD, Gayton D, Grad R, Huang A et al. Unnecessary prescribing of NSAIDs and the management of NSAID-related gastropathy in medical practice. *Ann Intern Med.* 1997; 127(6):429-438.

181. Federation of State Boards of Physical Therapy. *Job Analysis for U.S. Physical Therapy Practice.* 2002. Iowa City, Iowa, ACT, Inc.

182. Lee M, Moseley A, Refshauge K. Effect of feedback on learning a vertebral joint mobilization skill. *Phys Ther.* 1990; 70(2):97-102.

183. Keating J, Matyas TA, Bach TM. The effect of training on physical therapists' ability to apply specified forces of palpation. *Phys Ther.* 1993; 73(1):45-53.

184. Cohen E, Triano JJ, McGregor M, Papakyriakou M. Biomechanical performance of spinal manipulation therapy by newly trained vs. practicing providers: does experience transfer to unfamiliar procedures? *J Manipulative Physiol Ther.* 1995; 18(6):347-352.

185. Mennell J.M. *Back Pain.* Boston: Little Brown and Company, Inc., 1960.

186. Maitland GD. *Vertebral Manipulation.* 5th ed. Sydney, New South Wales, Australia: Butterworths, 1986.

187. Kaltenborn F. *The Spine: Basic Evaluation and Mobilization Techniques.* 3rd ed. Minneapolis.MN: Orthopaedic Physical Therapy Products, 1993.

188. Boyling JD, Palastanga N, Grieve GP. *Grieve's Modern Manual Therapy: The Vertebral Column*. 2nd ed. St. Louis, MO: Churchill Livingstone, 1994.

189. Greenman PE. *Principles of Manual Medicine*. 2nd ed. Philadelphia,P.A.: Lippincott Williams and Wilkins, 1996.

190. Winkel D, Aufdemkampe G, Matthijs O, Meijer OG, Phelps V. *Diagnosis and Treatment of the Spine*. Gaithersburg, MD: Aspen Publishers, Inc., 1996.

191. Paris SV. *Introduction to Spinal Evaluation and Manipulation: S1 Seminar Manual*. 3rd ed. St. Augustine, FL: Patris Press, 1997.

192. Flynn TW. Move it and move on. *J Orthop Sports Phys Ther*. 2002; 32(5):192-193.

193. Haas M, Groupp E, Panzer D, Partna L, Lumsden S, Aickin M. Efficacy of cervical endplay assessment as an indicator for spinal manipulation. *Spine*. 2003; 28(11):1091-1096.

194. Hayward RS, Wilson MC, Tunis SR, Bass EB, Guyatt G. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? The Evidence-Based Medicine Working Group. *JAMA*. 1995; 274(7):570-574.

195. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA*. 2000; 284(1):79-84.

196. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA*. 1997; 277(6):488-494.

197. Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med.* 1992; 21(4):384-390.

198. Stiell IG, Greenberg GH, Wells GA, McKnight RD, Cwinn AA, Cacciotti T et al. Derivation of a decision rule for the use of radiography in acute knee injuries. *Ann Emerg Med.* 1995; 26(4):405-413.

199. Stiell IG, Lesiuk H, Wells GA, McKnight RD, Brison R, Clement C et al. The Canadian CT Head Rule Study for patients with minor head injury: rationale, objectives, and methodology for phase I (derivation). *Ann Emerg Med.* 2001; 38(2):160-169.

200. Stiell IG, Lesiuk H, Wells GA, Coyle D, McKnight RD, Brison R et al. Canadian CT head rule study for patients with minor head injury: methodology for phase II (validation and economic analysis). *Ann Emerg Med.* 2001; 38(3):317-322.

201. Stiell IG, Wells GA, Vandemheen KL, Clement CM, Lesiuk H, De Maio VJ et al. The Canadian C-spine rule for radiography in alert and stable trauma patients. *JAMA.* 2001; 286(15):1841-1848.

202. Ebell MH, Smith MA, Barry HC, Ives K, Carey M. The rational clinical examination. Does this patient have strep throat? *JAMA.* 2000; 284(22):2912-2918.

203. Morise AP, Haddad WJ, Beckner D. Development and validation of a clinical score to estimate the probability of coronary artery disease in men and women presenting with suspected coronary disease. *Am J Med.* 1997; 102(4):350-356.

204. Ferreira G, Carson JL. Clinical prediction rules for the diagnosis of pulmonary embolism. *Am J Med.* 2002; 113(4):337-338.

205. van Walraven C, Forster AJ, Stiell IG. Derivation of a clinical decision rule for the discontinuation of in-hospital cardiac arrest resuscitations. *Arch Intern Med.* 1999; 159(2):129-134.

206. Gorelick MH, Shaw KN. Clinical decision rule to identify febrile young girls at risk for urinary tract infection. *Arch Pediatr Adolesc Med.* 2000; 154(4):386-390.

207. Sinclair DR, Chung F, Mezei G. Can postoperative nausea and vomiting be predicted? *Anesthesiology.* 1999; 91(1):109-118.

208. Hartling L, Pickett W, Brison RJ. Derivation of a clinical decision rule for whiplash associated disorders among individuals involved in rear-end collisions. *Accid Anal Prev.* 2002; 34(4):531-539.

209. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. Logistic Regression Analysis. *Applied Regression Analysis and Other Multivariable Methods.* Pacific Grove, CA: Duxbury Press, 1998: 656-686.

210. Fritz JM, Irrgang JJ. A Comparison of a Modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther.* 2001; 81(2):776-788.

211. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain.* 1994; 58(3):387-392.

212. Mann NH, Brown MD, Hertz DB, Enger I, Tompkins J. Initial-impression diagnosis using low-back pain patient pain drawings. *Spine*. 1993; 18(1):41-53.

213. Werneke M, Hart DL, Cook D. A descriptive study of the centralization phenomenon. A prospective analysis. *Spine*. 1999; 24(7):676-683.

214. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear- avoidance beliefs in chronic low back pain and disability. *Pain*. 1993; 52(2):157-168.

215. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy*. 1980; 66(8):271-273.

216. Stratford PW, Binkley J, Solomon P, Gill C, Finch E. Assessing change over time in patients with low back pain. *Phys Ther*. 1994; 74(6):528-533.

217. Fairbank J. Use of Oswestry Disability Index (ODI). *Spine*. 1995; 20(13):1535-1537.

218. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine*. 2000; 25(22):2940-2953.

219. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*. 1996; 65(1):71-76.

220. Waddell G, McCulloch JA, Kummel E, Venner RM. Nonorganic signs in low-back pain. *Spine*. 1980; 5:117-125.

221. Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. *Phys Ther*. 1998; 78(8):829-837.

222. Cibulka MT, Delitto A, Koldehoff RM. Changes in innominate tilt after manipulation of the sacroiliac joint in patients with low back pain. An experimental study. *Phys Ther.* 1988; 68(9):1359-1363.

223. Haas M, Peterson D, Rothman EH, Panzer D, Krein R, Johansen R et al. Responsiveness of leg alignment changes associated with articular pressure testing to spinal manipulation: the use of a randomized clinical trial design to evaluate a diagnostic test with a dichotomous outcome. *J Manipulative Physiol Ther.* 1993; 16(5):306-311.

224. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA.* 1995; 274(8):645-651.

225. Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain. Interrater reliability and short-term treatment outcomes. *Spine.* 2000; 25(1):106-114.

226. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM.* New York: Churchill Livingstone, Inc., 2000.

227. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA.* 1994; 271(9):703-707.

228. Fritz JM, Erhard RE, Hagen BF. Segmental instability of the lumbar spine. *Phys Ther.* 1998; 78(8):889-896.

229. Erhard RE. *The Spinal Exercise Handbook: A Home Exercise Manual for a Managed Care Environment.* Pittsburgh, PA: Laurel Concepts, 1998.

230. McGill SM. Low back exercises: Evidence for improving exercise regimens. *Phys Ther.* 1998; 78:754-764.

231. Battie MC, Cherkin DC, Dunn R, Ciol MA, Wheeler KJ. Managing low back pain: attitudes and treatment preferences of physical therapists. *Phys Ther.* 1994; 74(3):219-226.

232. Issacs ER, Bookhout MR, Woburn MA. *Bourdillon's Spinal Manipulation.* 6th ed. Oxford, United Kingdom: Butterworth-Heinemann Inc., 2002.

233. Evidence-Based Medicine Working Group. *Users' Guides to the Medical Literature.* 1st ed. Chicago: AMA Press, 2002.

234. Whitman JM, Flynn TW, Fritz JM, Wainner RS, Childs JD. The reliability of commonly used tests and measures in the examination and treatment of patients with low back pain. *J Orthop Sports Phys Ther.* In press.

235. Spitznagel EL, Helzer JE. A proposed solution to the base rate problem in the kappa statistic. *Arch Gen Psychiatry.* 1985; 42(7):725-728.

236. Chesworth BM, Padfield BJ, Helewa A. A comparison of hip mobility in patients with nonspecific low back pain. *Physiotherapy Canada.* 1994; 46:267-274.

237. Fairbank JC, Pynsent PB, Van Poortvliet JA, Phillips H. Influence of anthropometric factors and joint laxity in the incidence of adolescent back pain. *Spine.* 1984; 9(5):461-464.

238. Mellin G. Correlations of hip mobility with degree of back pain and lumbar spinal mobility in chronic low-back pain patients. *Spine*. 1988; 13(6):668-670.

239. Fritz JM, Whitman JM, Flynn TW, Wainner RS, Childs JD. Clinical factors related to the failure of individuals with low back pain to improve with spinal manipulation. *Physical Therapy*. In press.

240. Cibulka MT, Sinacore DR, Cromer GS, Delitto A. Unilateral hip rotation range of motion asymmetry in patients with sacroiliac joint regional pain. *Spine*. 1998; 23(9):1009-1015.

241. Cibulka MT. The treatment of the sacroiliac joint component to low back pain: a case report. *Phys Ther*. 1992; 72(12):917-922.

242. Ellison JB, Rose SJ, Sahrmann SA. Patterns of hip rotation range of motion: a comparison between healthy subjects and patients with low back pain. *Phys Ther*. 1990; 70(9):537-541.

243. McGinn T, Guyatt G, Wyer P, Naylor CD, Stiell I. Diagnosis: Clinical Prediction Rule. In: Guyatt G, Rennie D, editors. *Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. Chiacago: AMA Press, 2002: 471-483.

244. Fortin JD, Dwyer AP, West S, Pier J. Sacroiliac joint: pain referral maps upon applying a new injection/arthrography technique. Part I: Asymptomatic volunteers. *Spine*. 1994; 19(13):1475-1482.

245. Fritz JM, Delitto A, Welch WC, Erhard RE. Lumbar spinal stenosis: a review of current concepts in evaluation, management, and outcome measurements. *Arch Phys Med Rehabil*. 1998; 79(6):700-708.

246. Chan CW, Goldman S, Ilstrup DM, Kunselman AR, O'Neill PI. The pain drawing and Waddell's nonorganic physical signs in chronic low-back pain. *Spine.* 1993; 18(13):1717-1722.

247. Uden A, Astrom M, Bergenudd H. Pain drawings in chronic back pain. *Spine.* 1988; 13(4):389-392.

248. Werneke MW, Harris DE, Lichter RL. Clinical effectiveness of behavioral signs for screening chronic low-back pain patients in a work-oriented physical rehabilitation program. *Spine.* 1993; 18(16):2412-2418.

249. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. *Ann Rheum Dis.* 1978; 37(4):378-381.

250. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain.* 1986; 27(1):117-126.

251. Jensen MP, Miller L, Fisher LD. Assessment of pain during medical procedures: a comparison of three scales. *Clin J Pain.* 1998; 14(4):343-349.

252. Katz J, Melzack R. Measurement of pain. *Surg Clin North Am.* 1999; 79(2):231-252.

253. Jacob T, Baras M, Zeev A, Epstein L. Low back pain: reliability of a set of pain measurement tools. *Arch Phys Med Rehabil.* 2001; 82(6):735-742.

254. Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain.* 1999; 80(1-2):329-339.

255. Klenerman L, Slade PD, Stanley IM. The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. *Spine*. 1995; 20:478-484.

256. Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. *Pain*. 2001; 94(1):7-15.

257. Fritz JM, George SZ. Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs. *Phys Ther*. 2002; 82(10):973-983.

258. Glasziou P, Guyatt GH, Dans AL, Dans LF, Straus S, Sackett DL. Applying the results of trials and systematic reviews to individual patients [editorial] [see comments]. *ACP J Club*. 1998; 129(3):A15-A16.

259. Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine*. 1992; 17(6):617-628.

260. Hodges PW, Richardson CA. Inefficient muscular stabilization of the lumbar spine associated with low back pain. A motor control evaluation of transversus abdominis. *Spine*. 1996; 21(22):2640-2650.

261. Panjabi MM, Abumi K, Duranceau J, Oxland TR. Spinal stability and intersegmental muscle forces. A biomechanical model. *Spine*. 1989; 14(2):194-199.

262. Glass GV, Hopkins KD. *Statistical Methods in Education and Psychology*. 3rd ed. Boston: Allyn and Bacon, 1996.

263. Turk DC, Rudy TE, Sorkin BA. Neglected topics in chronic pain treatment outcome studies: determination of success. *Pain*. 1993; 53:3-16.

264. SPSS for Windows 10.1. Chicago, IL.: SPSS Inc., 1998.

265. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers, 1988.

266. Winer BJ, Brown DR, Michels KM. *Statistical Principles in Experimental Design.* 3rd ed. New York: McGraw-Hill, 1991.

267. Keppel G. *Design and Analysis: A Researcher's Handbook.* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1982.

268. Huynh H, Feldt LS. Estimation of the Box correction for degrees of freedom from sample data in the randomized block and split-plot designs. *J Educ Statist.* 1976; 1:69.

269. Geisser S, Greenhouse SW. An extension of Box's results on the use of the F distribution in mutivariate analysis. *Ann Math Statist.* 1958; 29:885.

270. Farrar JT, Young JP, Jr., LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain.* 2001; 94(2):149-158.

271. O'Sullivan PB, Phyty GD, Twomey LT, Allison GT. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. *Spine.* 1997; 22(24):2959-2967.

272. George SZ, Fritz JM, Bialosky JE, Donald DA. The effect of a fear-avoidance based physical therapy intervention for patients with acute low back pain: Results of a randomized clinical trial. *Spine.* In press.

273. Hicks GE, Fritz JM, Delitto A, McGill SM. Predictive validity of clinical variables used in the determination of patient prognosis following a lumbar stabilization program. *Spine*. In press.

274. Federation of State Boards of Physical Therapy. *The Federation of State Boards of Physical Therapy website*. 2003. Federation of State Boards of Physical Therapy.

275. Bach E. Job Analysis. *Federation Forum*. 2002; 17(3):6-7.

276. Federation of State Boards of Physical Therapy. *Physical Therapist Content Outline*. 2002. Alexandria, VA.

277. Haug J, Wood LT. Efficacy of neuromuscular stimulation of the quadriceps femoris during continuous passive motion following total knee arthroplasty. *Arch Phys Med Rehabil*. 1988; 69(6):423-424.

278. Faghri PD, Rodgers MM, Glaser RM, Bors JG, Ho C, Akuthota P. The effects of functional electrical stimulation on shoulder subluxation, arm function recovery, and shoulder pain in hemiplegic stroke patients. *Arch Phys Med Rehabil*. 1994; 75(1):73-79.

279. Snyder-Mackler L, Delitto A, Bailey SL, Stralka SW. Strength of the quadriceps femoris muscle and functional recovery after reconstruction of the anterior cruciate ligament. A prospective, randomized clinical trial of electrical stimulation. *J Bone Joint Surg Am*. 1995; 77(8):1166-1173.

280. Chantraine A, Baribeault A, Uebelhart D, Gremion G. Shoulder pain and dysfunction in hemiplegia: effects of functional electrical stimulation. *Arch Phys Med Rehabil*. 1999; 80(3):328-331.

281. Brosseau LU, Pelland LU, Casimiro LY, Robinson VI, Tugwell PE, Wells GE. Electrical stimulation for the treatment of rheumatoid arthritis. *Cochrane Database Syst Rev.* 2002;(2):CD003687.

282. Gotlin RS, Hershkowitz S, Juris PM, Gonzalez EG, Scott WN, Insall JN. Electrical stimulation effect on extensor lag and length of hospital stay after total knee arthroplasty. *Arch Phys Med Rehabil.* 1994; 75(9):957-959.

283. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for low back pain. *Phys Ther.* 2001; 81(10):1641-1674.

284. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for neck pain. *Phys Ther.* 2001; 81(10):1701-1717.

285. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for shoulder pain. *Phys Ther.* 2001; 81(10):1719-1730.

286. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for knee pain. *Phys Ther.* 2001; 81(10):1675-1700.

287. Hurwitz EL, Morgenstern H, Harber P, Kominski GF, Belin TR, Yu F et al. Second Prize: The effectiveness of physical modalities among patients with low back pain randomized to chiropractic care: findings from the UCLA low back pain study. *J Manipulative Physiol Ther.* 2002; 25(1):10-20.

288. van der Windt DA.van der Heijden GJ.van den Berg SG.ter Riet. Ultrasound therapy for musculoskeletal disorders: a systematic review. *Pain.* 1999; 81(3):257-271.

289. Nadler SF, Steiner DJ, Erasala GN, Hengehold DA, Hinkle RT, Beth GM et al. Continuous low-level heat wrap therapy provides more efficacy than Ibuprofen and acetaminophen for acute low back pain. *Spine*. 2002; 27(10):1012-1017.

290. Nadler SF, Steiner DJ, Petty SR, Erasala GN, Hengehold DA, Weingand KW. Overnight use of continuous low-level heatwrap therapy for relief of low back pain. *Arch Phys Med Rehabil*. 2003; 84(3):335-342.

291. Pages IH, Jahr S, Schaufele MK, Conradi E. Comparative analysis of biofeedback and physical therapy for treatment of urinary stress incontinence in women. *Am J Phys Med Rehabil*. 2001; 80(7):494-502.

292. Aukee P, Immonen P, Penttinen J, Laippala P, Airaksinen O. Increase in pelvic floor muscle activity after 12 weeks' training: a randomized prospective pilot study. *Urology*. 2002; 60(6):1020-1023.

293. Burgio KL, Locher JL, Goode PS, Hardin JM, McDowell BJ, Dombrowski M et al. Behavioral vs drug treatment for urge urinary incontinence in older women: a randomized controlled trial. *JAMA*. 1998; 280(23):1995-2000.

294. Morkved S, Bo K, Fjortoft T. Effect of adding biofeedback to pelvic floor muscle training to treat urodynamic stress incontinence. *Obstet Gynecol*. 2002; 100(4):730-739.

295. Burgio KL, Goode PS, Locher JL, Umlauf MG, Roth DL, Richter HE et al. Behavioral training with and without biofeedback in the treatment of urge incontinence in older women: a randomized controlled trial. *JAMA*. 2002; 288(18):2293-2299.

296. Hay-Smith EJ, Bo Berghmans LC, Hendriks HJ, de Bie RA, van Waalwijk van Doorn ES. Pelvic floor muscle training for urinary incontinence in women. *Cochrane Database Syst Rev.* 2001;(1):CD001407.

297. Osiri M, Welch V, Brosseau L, Shea B, McGowan J, Tugwell P et al. Transcutaneous electrical nerve stimulation for knee osteoarthritis. *Cochrane Database Syst Rev.* 2000;(4):CD002823.

298. Carroll D, Moore RA, McQuay HJ, Fairman F, Tramer M, Leijon G. Transcutaneous electrical nerve stimulation (TENS) for chronic pain. *Cochrane Database Syst Rev.* 2001;(3):CD003222.

299. Milne S, Welch V, Brosseau L, Saginur M, Shea B, Tugwell P et al. Transcutaneous electrical nerve stimulation (TENS) for chronic low back pain. *Cochrane Database Syst Rev.* 2001;(2):CD003008.

300. Davies DM, Johnston DW, Beaupre LA, Lier DA. Effect of adjunctive range-of-motion therapy after primary total knee arthroplasty on the use of health services after hospital discharge. *Can J Surg.* 2003; 46(1):30-36.

301. Beaupre LA, Davies DM, Jones CA, Cinats JG. Exercise combined with continuous passive motion or slider board therapy compared with exercise only: a randomized controlled trial of patients following total knee arthroplasty. *Phys Ther.* 2001; 81(4):1029-1037.

302. Lau SK, Chiu KY. Use of continuous passive motion after total knee arthroplasty. *J Arthroplasty.* 2001; 16(3):336-339.

303. Chen B, Zimmerman JR, Soulen L, DeLisa JA. Continuous passive motion after total knee arthroplasty: a prospective study. *Am J Phys Med Rehabil.* 2000; 79(5):421-426.

304. MacDonald SJ, Bourne RB, Rorabeck CH, McCalden RW, Kramer J, Vaz M. Prospective randomized clinical trial of continuous passive motion after total knee arthroplasty. *Clin Orthop.* 2000;(380):30-35.

305. Brosseau L, Casimiro L, Milne S, Robinson V, Shea B, Tugwell P et al. Deep transverse friction massage for treating tendinitis. *Cochrane Database Syst Rev.* 2002;(4):CD003528.

306. Furlan AD, Brosseau L, Imamura M, Irvin E. Massage for low back pain. *Cochrane Database Syst Rev.* 2002;(2):CD001929.

307. Stephens EB. Manipulative therapy in physical therapy curricula. *Phys Ther.* 1973; 53(1):40-50.

308. Ben Sorek S, Davis CM. Joint mobilization education and clinical use in the United States. *Phys Ther.* 1988; 68(6):1000-1004.

309. Bryan JM, McClune LD, Romito S, Stetts DM, Finstuen K. Spinal mobilization curricula in professional physical therapy education programs. *Journal of Physical Therapy Education.* 1997; 11(2):11-15.

310. Boissonnault W, Bryan JM, Fox KJ. Joint manipulation curricula in first-professional physical therapy education programs. *J Orthop Sports Phys Ther.* In press.

311. Shekelle PG, Adams AH, Chassin MR. *The appropriateness of spinal manipulation for low back pain: Project overview and literature review.* 1991. Santa Monica, CA, RAND Corporation.

312. American Physical Therapy Association (APTA). *A Normative Model of Physical Therapist Professional Education. Version 2000.* Alexandria, VA: APTA Education Division, 2000.

313. Commission on Accreditation of Physical Therapy Education. *Evaluative Criteria for Accreditation of Education Programs for the Preparation of Physical Therapists.* Alexandria, VA: American Physical Therapy Association, 1998.

314. Albert CM, Mittleman MA, Chae CU, Lee IM, Hennekens CH, Manson JE. Triggering of sudden death from cardiac causes by vigorous exertion. *N Engl J Med.* 2000; 343(19):1355-1361.

315. Eck JC, Circolone NJ. The use of spinal manipulation in the treatment of low back pain: a review of goals, patient selection, techniques, and risks. *J Orthop Sci.* 2000; 5(4):411-417.

316. Ellison JB, Rose SJ, Sahrmann SA. Patterns of hip rotation range of motion: a comparison between healthy subjects and patients with low back pain. *Phys Ther.* 1990; 70(9):537-541.

317. Cibulka MT, Delitto A. A comparison of two different methods to treat hip pain in runners. *J Orthop Sports Phys Ther.* 1993; 17(4):172-176.

318. Cibulka MT. Low back pain and its relation to the hip and foot. *J Orthop Sports Phys Ther.* 1999; 29(10):595-601.

319. Auleley GR, Ravaud P, Giraudeau B, Kerboull L, Nizard R, Massin P et al. Implementation of the Ottawa ankle rules in France. A multicenter randomized controlled trial. *JAMA*. 1997; 277(24):1935-1939.

320. Stiell IG, McKnight RD, Greenberg GH, McDowell I, Nair RC, Wells GA et al. Implementation of the Ottawa ankle rules. *JAMA*. 1994; 271(11):827-832.

321. Verbeek PR, Stiell IG, Hebert G, Sellens C. Ankle radiograph utilization after learning a decision rule: a 12-month follow-up. *Acad Emerg Med.* 1997; 4(8):776-779.

322. Stiell IG, Wells GA, Hoag RH, Sivilotti ML, Cacciotti TF, Verbeek PR et al. Implementation of the Ottawa Knee Rule for the use of radiography in acute knee injuries. *JAMA*. 1997; 278(23):2075-2079.

323. Nichol G, Stiell IG, Wells GA, Juergensen LS, Laupacis A. An economic analysis of the Ottawa knee rule. *Ann Emerg Med.* 1999; 34(4 Pt 1):438-447.

324. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ.* 1995; 153(10):1423-1431.

325. Cameron C, Naylor CD. No impact from active dissemination of the Ottawa Ankle Rules: further evidence of the need for local implementation of practice guidelines. *CMAJ.* 1999; 160(8):1165-1168.

326. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. *JAMA*. 1995; 274(9):700-705.

327. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA*. 1999; 282(15):1458-1465.

328. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA*. 1999; 282(9):867-874.

329. Ehrmann-Feldman D, Rossignol M, Abenhaim L, Gobeille D. Physician referral to physical therapy in a cohort of workers compensated for low back pain. *Phys Ther*. 1996; 76(2):150-156.

330. Curtis P, Carey TS, Evans P, Rowane MP, Mills GJ, Jackman A. Training primary care physicians to give limited manual therapy for low back pain: patient outcomes. *Spine*. 2000; 25(22):2954-2960.